Dynamic Systems and Applications 26 (2017) 627-662

CONSTRAINED LINEAR-QUADRATIC CONTROL PROBLEMS OVER TIME SCALES AND WEAK NORMALITY

VERA ZEIDAN

Department of Mathematics, Michigan State University
East Lansing, MI 48824-1027, USA
zeidan@math.msu.edu

ABSTRACT. Time scale linear-quadratic control problems with affine mixed state-control and joint endpoints equality constraints are considered. Without any controllability assumption, it is shown that the feasible pairs at which the first variation vanishes are exactly the feasible pairs that satisfy the weak maximum principle (called "extremals") with $\lambda_0 = 1$. In this case, we say the problem is "weakly normal" at such feasible pairs. When a certain matrix function $S(\cdot)$ has invertible images, the weak-normality condition at $(\bar{x}, \bar{u})$ with associated adjoint variable $\bar{p}$ is also equivalent to $(\bar{x}, \bar{p})$ solving the corresponding *non-homogeneous symplectic boundary value problem* and to $\bar{u}$ being a certain affine combination of this solution. In this equivalence, the invertibility of the corresponding matrices $S(t)$ is not needed when the linear-quadratic problem is itself *symplectic*. As an application, it is established that without any controllability assumption, the optimality in linear-quadratic problems is characterized in terms of either the weak normality condition, or the solvability of the corresponding symplectic boundary value. These results are obtained for linear-quadratic control problems with or without shift in the state variable and are new not only for the time scale setting but also for the continuous time and discrete time settings.

Keywords. Time scale; Linear-Quadratic control problems; Weak maximum principle; Weak-normality; Symplectic systems; Symplectic quadratic problems; Nonlinear optimal control; Second variation

AMS (MOS) Subject Classification. 49K15, 39A12, 34K35, 49N25, 93B05

1. INTRODUCTION

Consider the *linear-quadratic optimal control problem* on a time scale $\mathbb{T}$ having constraints as variable state endpoints equalities as well as mixed control–state equalities. As is customary of dynamic and variational problems over time scales, there are two formulations: one we denote by $(\mathcal{LQ}^\sigma)$, in which the data has a *shift* in the state variable x^σ, and the second is $(\mathcal{LQ})$, in which x has *no shift*. When the time scale $\mathbb{T}$ is a connected interval both forms collapse to the continuous time linear-quadratic control problem. The presence of two forms of such variational problems stems from the fact that the time scale incorporates also the discrete setting where both forms are prominent; see e.g. [7, 27] for the discrete problems with *no shift*, and [3, 16, 18, 26] for

Received June 3, 2017

the discrete problems with *shift*. The framework of time scale dynamic equation and its corresponding objective functional includes not only the continuous and discrete time setting but also anything in between see e.g., [1, 2, 8, 10, 11, 17].

Let $\mathbf{P}(\cdot)$, $\mathbf{Q}(\cdot)$, $\mathbf{R}(\cdot)$, $\mathbf{A}(\cdot)$, $\mathbf{B}(\cdot)$, $\mathbf{K}(\cdot)$, $\mathbf{N}(\cdot)$, $\mathbf{c}(\cdot)$, $\mathbf{z}(\cdot)$ and $\mathbf{k}(\cdot)$ be respectively $n\times n$-, $n\times m$-, $m\times m$-, $n\times n$-, $n\times m$-, $k\times n$-, $k\times m$-, $k\times 1$-, $n\times 1$-, and $m\times 1$- matrix-functions defined over the time scale $[a,\rho(b)]_{\mathbb{T}}$ which are piecewise rd-continuous with $\mathbf{P}(t)$ and $\mathbf{R}(t)$ symmetric for all t, $m\le n$, and $k\le m$. Let $\mathbf{\Gamma}$ and $\mathbf{M}$ be respectively $2n\times 2n$- and $r\times 2n$-matrices with $\mathbf{\Gamma}$ symmetric and $r\le 2n$, and let $\mathbf{d}$ be a vector in $\mathbb{R}^r$ and $\mathbf{c_a}$ and $\mathbf{c_b}$ be vectors in $\mathbb{R}^n$. Assume in addition that $\mathbf{M}$ and $\mathbf{N}(t)$, for all t, are of full rank. Define

$$\begin{aligned}\mathbf{F}(x,u) &:= \mathbf{c_a}^T x(a) + \mathbf{c_b}^T x(b) + \frac{1}{2}\begin{pmatrix} x(a)\\ x(b)\end{pmatrix}^T \mathbf{\Gamma}\begin{pmatrix} x(a)\\ x(b)\end{pmatrix}\\ &\quad + \int_a^b \Big[\mathbf{z}^T x^\sigma + \mathbf{k}^T u + \frac{1}{2}\Big((x^\sigma)^T\mathbf{P}x^\sigma + 2(x^\sigma)^T\mathbf{Q}\,u + u^T\mathbf{R}u\Big)\Big](t)\Delta t. \qquad (1.1)\end{aligned}$$

The linear-quadratic control problem *with* shift in x, $(\mathcal{LQ}^\sigma)$, is defined as:

$(\mathcal{LQ}^\sigma)$ minimize $\mathbf{F}(x,u)$

subject to: $x\in \mathrm{C}^1_{\mathrm{prd}}[a,b]_{\mathbb{T}}$ and $u\in \mathrm{C}_{\mathrm{prd}}[a,\rho(b)]_{\mathbb{T}}$ satisfying

$$x^\Delta(t) = \mathbf{A}(t)x^\sigma(t) + \mathbf{B}(t)u(t), \quad t\in[a,\rho(b)]_{\mathbb{T}}, \qquad (1.2)$$

$$\mathbf{K}(t)x^\sigma(t) + \mathbf{N}(t)u(t) = \mathbf{c}(t), \quad t\in[a,\rho(b)]_{\mathbb{T}}, \qquad (1.3)$$

$$\mathbf{M}\begin{pmatrix} x(a)\\ x(b)\end{pmatrix} = \mathbf{d}, \qquad (1.4)$$

where $\mathbb{T}$ is a bounded time scale, $a:=\min\mathbb{T}$, $b:=\max\mathbb{T}$, and $[c,d]_{\mathbb{T}} := [c,d]\cap\mathbb{T}$. The state $x:[a,b]_{\mathbb{T}}\to\mathbb{R}^n$ is piecewise rd-continuously Δ-differentiable, and the control $u:[a,\rho(b)]_{\mathbb{T}}\to\mathbb{R}^m$ is piecewise rd-continuous. The Hamiltonian corresponding to problem $(\mathcal{LQ}^\sigma)$ is

$$\begin{aligned}\mathbf{H}(t,x,u,p,\lambda,\lambda_0) &:= p^T[\mathbf{A}(t)x+\mathbf{B}(t)u] + \lambda^T[\mathbf{K}(t)x+\mathbf{N}(t)u-\mathbf{c}(t)] + \lambda_0\Big[\mathbf{z}^T(t)x\\ &\quad + \mathbf{k}^T(t)u + \frac{1}{2}\Big(x^T\mathbf{P}(t)x + 2x^T\mathbf{Q}(t)u + u^T\mathbf{R}(t)u\Big)\Big]. \qquad (1.5)\end{aligned}$$

A pair (x,u) satisfying (1.2), (1.3) and (1.4) is said to be *feasible* for $(\mathcal{LQ}^\sigma)$. A feasible pair $(\bar{x},\bar{u})$ is a *weak local minimum* for $(\mathcal{LQ}^\sigma)$ if there exists $\bar{\varepsilon}>0$ such that for any feasible (x,u) with $\|x-\bar{x}\|_{\mathrm{C}}<\bar{\varepsilon}$ and $\|u-\bar{u}\|_{\mathrm{C}_{\mathrm{prd}}}<\bar{\varepsilon}$ we have $\mathbf{F}(\bar{x},\bar{u})\le\mathbf{F}(x,u)$, where

$$\|u\|_{\mathrm{C}_{\mathrm{prd}}} := \sup_{t\in[a,\rho(b)]_{\mathbb{T}}} |u(t)|, \qquad \|x\|_{\mathrm{C}} := \max_{t\in[a,b]_{\mathbb{T}}} |x(t)|.$$

Given the data $P(\cdot)$, $Q(\cdot)$, $R(\cdot)$, $A(\cdot)$, $B(\cdot)$, $K(\cdot)$, $N(\cdot)$, $c(\cdot)$, $z(\cdot)$, $k(\cdot)$, Γ, M, d, c_a, and c_b, satisfying the assumptions of the corresponding coefficients in $(\mathcal{LQ}^\sigma)$. Set

$$F(x,u) := {c_a}^T x(a) + {c_b}^T x(b) + \frac{1}{2}\begin{pmatrix} x(a) \\ x(b) \end{pmatrix}^T \Gamma \begin{pmatrix} x(a) \\ x(b) \end{pmatrix}$$

$$+ \int_a^b \Big[z^T x + k^T u + \frac{1}{2}\big(x^T P x + 2x^T Q\, u + u^T R u\big)\Big](t)\Delta t. \tag{1.6}$$

Then, the linear-quadratic control problem with *no* shift in x, $(\mathcal{LQ})$, is defined as

$(\mathcal{LQ})$ minimize $\quad F(x,u)$

subject to:$x \in \mathrm{C}^1_{\mathrm{prd}}[a,b]_{\mathbb{T}}$ and $u \in \mathrm{C}_{\mathrm{prd}}[a,\rho(b)]_{\mathbb{T}}$ satisfying

$$x^\Delta(t) = A(t)x(t) + B(t)u(t), \quad t \in [a,\rho(b)]_{\mathbb{T}}, \tag{1.7}$$

$$K(t)x(t) + N(t)u(t) = c(t), \quad t \in [a,\rho(b)]_{\mathbb{T}}, \tag{1.8}$$

$$M\begin{pmatrix} x(a) \\ x(b) \end{pmatrix} = d, \tag{1.9}$$

where the state x, the control u, the Hamiltonian H, and the notion of *weak local* minimality are defined in terms of the data of this problem similarly to those for the problem $(\mathcal{LQ}^\sigma)$. The *feasibility* of a pair (x,u) means that it satisfies (1.7)–(1.9).

Note that the time scale $[a,b]_{\mathbb{T}}$ does not only incorporate both the continuous time connected interval $[a,b]$ and the discrete time interval $[0,N+1]_{\mathbb{N}} := \{0,1,\dots,N+1\}$ but also, as is the case in many applications [4, 6, 19, 20], the time scale encompasses the setting where the time is *neither* continuous *nor* discrete. Thorough description of the time scales theory can be found in [6]. The functions σ and ρ are respectively the forward and backward jump operators on $[a,b]_{\mathbb{T}}$ and $x^\sigma(t) := x\big(\sigma(t)\big)$. The time scale Δ-derivative and the corresponding integral are denoted, respectively, by $x^\Delta(t)$ and $\int_a^b G(t)\,\Delta t$. In the special cases of the continuous and discrete times, $x^\Delta(t)$ reduces to the standard derivative $\dot{x}(t)$ and forward difference $\Delta x(t)$, the integral to $\int_a^b G(t)\,dt$ and $\sum_{k=0}^N G(k)$, and the jump operators to $\sigma(t) = t = \rho(t)$ and $\sigma(k) = k+1$, $\rho(k) = k-1$, respectively. The graininess function on $[a,b]_{\mathbb{T}}$ is $\mu(t) := \sigma(t) - t$. The notions of piecewise rd-continuous ($\mathrm{C}_{\mathrm{prd}}$) and piecewise rd-continuously Δ-differentiable ($\mathrm{C}^1_{\mathrm{prd}}$) functions are introduced in [17].

In the presence of variable endpoints constraints and *only* pure-control equality constraints, first and second-order order necessary conditions for weak local optimality invoking *Lagrange multipliers* were derived in [23] for the time scale *nonlinear* optimal control problem and are known as the weak maximum principle and the second variation condition, respectively. This latter states that the linear-quadratic *accessory* problem corresponding to the nonlinear problem has zero as a minimum value. These necessary conditions unify the corresponding continuous time results, see e.g. [12], and the discrete time results, see e.g. [15] or [16]. They also extend the results in [5, 26] from the calculus of variations to general optimal control problems. In order that these conditions be meaningful or *non-degenerate*, the multiplier λ_0 associated to the objective function must be non-zero. A sufficient condition for the non-degeneracy is the *M-controllability* of the constraints linearization (see [23]).

Optimal control problems with *mixed control-state* constraints were never studied over time scales. However, they were extensively studied in the*continuous* time setting for which such first and second order optimality conditions are obtained (see e.g, [29], [30], [28]). In order to have $\lambda_0 = 1$, it is customary to either assume one state endpoint is free (e.g., [9]), which yields the controllability, or to directly assume the controllability of the constraints linearization when both endpoints vary (e.g., [34]). On the other hand, in the continuous time setting, when the problem is *purely linear-quadratic* (see Remark 2.2) over $L^2[a, b]$ controls, the state endpoints are both fixed, and no mixed control-state constraints are present, it is shown in [13, Section 8] by means of Hilbert space methods and without any controllability, that any optimal solution must have at least one set of multipliers for which $\lambda_0 = 1$.

One main goal of this paper is to show without any controllability assumption that *any* linear-quadratic problem on time scales of the form $(\mathcal{LQ}^\sigma)$ or $(\mathcal{LQ})$ has the property that its "critical pairs" are exactly "extremals" at which the problem is "weakly normal." In other words, feasible pairs $(\bar{x}, \bar{u})$ at which the first variation of the problem vanishes are exactly those who admit at least one set of multipliers satisfying the weak maximum principle with $\lambda_0 = 1$. Once this result is established, a characterization of optimality for linear-quadratic problems is obtained in terms of the non-negativity of the second variation and the weak normality condition. Furthermore, the normality condition is shown to be also equivalent to the existence of a solution to the corresponding *non-homogeneous symplectic system* with boundary conditions when a certain matrix function $\mathbf{S}(\cdot)$ for $(\mathcal{LQ}^\sigma)$ or $S(\cdot)$ for $(\mathcal{LQ})$ have invertible images. This equivalence, however, does not require the invertibility of the corresponding matrix $S(t)$ when the linear-quadratic problem itself is *symplectic*; see Section 4 for details. In order to prove these results for the linear-quadratic problem in the form of $(\mathcal{LQ})$, it is shown that this latter is in fact equivalent to the form $(\mathcal{LQ}^\sigma)$; a result that is important on its own for this general setting.

The results in this paper are new, not only for the time scale setting, but also for the continuous or discrete time setting. In particular, the result in [13] for the continuous time is now extended to the case when the mixed state-control and the joint endpoints constraints are present and when the quadratic function is not necessarily *purely* quadratic (see Remark 2.2). Furthermore, the same result is now valid when the control space is the smaller set of piecewise continuous functions.

2. LINEAR-QUADRATIC PROBLEMS WITH STATE SHIFT

In the first part of this section we derive a characterization of optimality for the linear-quadratic control problem $(\mathcal{LQ}^\sigma)$ introduced in Section 1 without any $\mathbf{M}$-controllability assumption. This characterization is done in terms of the first variation being zero; a condition shown later to be equivalent to the "weak-normality" of the problem $(\mathcal{LQ}^\sigma)$ at the extremal.

Definition 2.1. A pair (η, v) is said to be an *admissible direction* for the problem $(\mathcal{LQ}^\sigma)$ if it satisfies the equation of motion (1.2), and the linearization of (1.3) and (1.4), that is,

$$\eta^\Delta(t) = \mathbf{A}(t)\eta^\sigma(t) + \mathbf{B}(t)v(t), \quad t \in [a, \rho(b)]_{\mathbb{T}}, \tag{2.1}$$

$$\mathbf{K}(t)\eta^\sigma(t) + \mathbf{N}(t)v(t) = 0, \quad t \in [a, \rho(b)]_{\mathbb{T}}, \tag{2.2}$$

$$\mathbf{M}\begin{pmatrix}\eta(a)\\ \eta(b)\end{pmatrix} = 0. \tag{2.3}$$

Note that when $\mathbf{c} \equiv 0$ and $\mathbf{d} = 0$ then feasible pairs and admissible directions are the same.

Since $\mathbf{N}(t)$ is of full rank for all $t \in [a, \rho(b)]_{\mathbb{T}}$, choose a function $\mathbf{Y} : [a, \rho(b)]_{\mathbb{T}} \to \mathbb{R}^{m\times(m-k)}$, $\mathbf{Y} \in \mathrm{C}_{\mathrm{prd}}$, such that, for each t, $\mathbf{Y}(t)$ is a matrix whose columns form an orthonormal basis for the space $\operatorname{Ker}\mathbf{N}(t)$. Set

$$\mathbf{N}_\dagger(t) := \mathbf{N}^T(t)(\mathbf{N}(t)\mathbf{N}^T(t))^{-1} \quad \text{for all } t \in [a, \rho(b)]_{\mathbb{T}}. \tag{2.4}$$

Then, equation (2.2) is equivalent to

$$v(t) = \mathbf{Y}(t)w(t) - \mathbf{N}_\dagger(t)\mathbf{K}(t)\eta^\sigma(t), \quad \text{for all } t \in [a, \rho(b)]_{\mathbb{T}}, \tag{2.5}$$

where $w(\cdot)$ is in $\mathrm{C}_{\mathrm{prd}}$. Hence, (η, v) satisfies Definition (2.1) is equivalent to v being defined through w via (2.5) and (η, w) satisfies

$$\eta^\Delta(t) = \Big(\mathbf{A}(t) - \mathbf{B}(t)\mathbf{N}_\dagger(t)\mathbf{K}(t)\Big)\eta^\sigma(t) + \mathbf{B}(t)\mathbf{Y}(t)w(t), \quad t \in [a, \rho(b)]_{\mathbb{T}}, \tag{2.6}$$

$$\mathbf{M}\begin{pmatrix}\eta(a)\\ \eta(b)\end{pmatrix} = 0. \tag{2.7}$$

We define the first and second variations of the problem $(\mathcal{LQ}^\sigma)$ at a pair $(\bar{x}, \bar{u})$ in $\mathrm{C}^1_{\mathrm{prd}} \times \mathrm{C}_{\mathrm{prd}}$ in the direction of an admissible pair (η, v) by

$$\begin{aligned}\mathbf{F}'(\bar{x}, \bar{u}; \eta, v) := {\mathbf{c_a}}^T\eta(a) + {\mathbf{c_b}}^T\eta(b) + \begin{pmatrix}\bar{x}(a)\\ \bar{x}(b)\end{pmatrix}^T \mathbf{\Gamma}\begin{pmatrix}\eta(a)\\ \eta(b)\end{pmatrix} + \int_a^b \Big[\mathbf{z}^T\eta^\sigma + \mathbf{k}^T v \\ + (\bar{x}^\sigma)^T\mathbf{P}\eta^\sigma + \bar{u}^T\mathbf{Q}^T\eta^\sigma + (\bar{x}^\sigma)^T\mathbf{Q}v + \bar{u}^T\mathbf{R}v\Big](t)\Delta t,\end{aligned} \tag{$\mathcal{L}^\sigma$}$$

$$(\mathcal{Q}^\sigma) \quad \mathbf{F}''(\bar{x},\bar{u};\eta,v) := \begin{pmatrix}\eta(a)\\ \eta(b)\end{pmatrix}^T \mathbf{\Gamma} \begin{pmatrix}\eta(a)\\ \eta(b)\end{pmatrix} + \int_a^b [(\eta^\sigma)^T \mathbf{P}\eta^\sigma + 2(\eta^\sigma)^T \mathbf{Q} v + v^T \mathbf{R} v](t)\Delta t.$$

Remark 2.2. (i) The affine feature of (1.2)–(1.4) and the quadraticity of the functional $\mathbf{F}$ in (1.1) yield that the second variation $\mathbf{F}''$ is independent of $(\bar{x},\bar{u})$.

(ii) When the problem is *purely* linear-quadratic, that is, $\mathbf{c_a} = 0$, $\mathbf{c_b} = 0$, $\mathbf{c} \equiv 0$, $\mathbf{z} \equiv 0$, $\mathbf{k} \equiv 0$, and $\mathbf{d} = 0$ then, feasible directions are also admissible directions and we readily obtain from $(\mathcal{L}^\sigma)$ that $\mathbf{F}'(\bar{x},\bar{u};\bar{x},\bar{u}) = 2\mathbf{F}(\bar{x},\bar{u})$ and from $(\mathcal{Q}^\sigma)$ that $\mathbf{F}''(\bar{x},\bar{u};\eta,v) = 2\mathbf{F}(\eta,v)$.

When $\mathbf{K} \equiv 0$, our problem $(\mathcal{LQ}^\sigma)$ is a special case of the nonlinear control problem studied in [23] for *pure* equality control constraints. A necessary condition for optimality is obtained in [23, Corollary 6.3]. It states that the first variation at an optimal pair must be equal *zero* along any admissible direction (η,v), provided $I - \mu(t)\mathbf{A}(t)$ is *invertible* for all t and the system (1.2) is $\mathbf{M}$-*controllable* over the null space of $\mathbf{N}$ (see Definition 4.2 therein, or take $\mathbf{K} \equiv 0$ in the $\mathbf{M}$-controllability notion given by (2.10)). Also, these same assumptions are required in [23, Theorem 7.2] for the second variation to be non-negative along any admissible direction (η,v).

Unlike the results in [23, Corollary 6.3 and Theorem 7.2] for the case when $\mathbf{K} \equiv 0$, the following proposition shows that even in the presence of *mixed* state-control constraints in the linear-quadratic problem $(\mathcal{LQ}^\sigma)$, the first and second order necessary conditions for $(\mathcal{LQ}^\sigma)$ do not require neither the invertibility of $I - \mu(\mathbf{A} - \mathbf{B}\mathbf{N}_\dagger\mathbf{K})$ nor the $\mathbf{M}$-controllability of (2.1)–(2.2). In addition, this Proposition shows that these necessary conditions are also sufficient for the weak local optimality of $(\bar{x},\bar{u})$, which is equivalent to the global optimality.

Proposition 2.3 (Optimality & 1st and 2nd variations of $(\mathcal{LQ}^\sigma)$)**.** *Let $(\bar{x},\bar{u})$ be feasible for $(\mathcal{LQ}^\sigma)$. Then, the following conditions are equivalent.*

(a) $(\bar{x},\bar{u})$ *is a weak local minimum for* $(\mathcal{LQ}^\sigma)$.
(b) *For all admissible directions* (η,v), $\mathbf{F}'(\bar{x},\bar{u};\eta,v) = 0$ *and* $\mathbf{F}''(\bar{x},\bar{u};\eta,v) \geq 0$.
(c) $(\bar{x},\bar{u})$ *is a global minimum for* $(\mathcal{LQ}^\sigma)$.

Remark 2.4. When($\mathcal{LQ}^\sigma$) is *purely* linear-quadratic and $(\bar{x},\bar{u})$ is a weak local minimum, then (a) $\Rightarrow$ (b) here and part (ii) of Remark 2.2 imply that $\mathbf{F}(\bar{x},\bar{u}) = 0$.

Proof of Proposition 2.3.

$(a) \Rightarrow (b)$: Let (η,v) be an admissible direction. Then, for any ε, the pair $(\bar{x}+\varepsilon\eta, \bar{u}+\varepsilon v)$ is feasible for problem $(\mathcal{LQ}^\sigma)$. By the weak local minimality of $(\bar{x},\bar{u})$, there exists $\varepsilon_0 > 0$ such that

$$\mathbf{F}(\bar{x}+\varepsilon\eta, \bar{u}+\varepsilon v) - \mathbf{F}(\bar{x},\bar{u}) \geq 0 \quad \forall \varepsilon \in [0,\varepsilon_0]. \tag{2.8}$$

Use (1.1) and foil $\mathbf{F}(\bar{x}+\varepsilon\eta, \bar{u}+\varepsilon v)$, it follows from $(\mathcal{L}^\sigma)$ that (2.8) is equivalent to:

$$\varepsilon^2\mathbf{F}''(\bar{x},\bar{u};\eta,v) + 2\varepsilon\mathbf{F}'(\bar{x},\bar{u};\eta,v) \geq 0 \quad \forall\varepsilon \in [0,\varepsilon_0]. \tag{2.9}$$

This implies that for all ε in $(0,\varepsilon_0]$, $\varepsilon\mathbf{F}''(\bar{x},\bar{u};\eta,v) + 2\,\mathbf{F}'(\bar{\eta},\bar{v};\eta,v) \geq 0$. Upon taking $\varepsilon \searrow 0$ it follows that

$$\mathbf{F}'(\bar{\eta},\bar{v};\eta,v) \geq 0 \quad \forall \text{ admissible directions } (\eta,v).$$

Since the set of admissible pairs is symmetric with respect to the origin and $\mathbf{F}'$ is linear in (η,v), it follows that $\mathbf{F}'(\bar{x},\bar{u};\eta,v) = 0$ for all admissible directions (η,v). By using $\mathbf{F}'(\bar{x},\bar{u};\eta,v) = 0$ in (2.9) it follows that $\mathbf{F}''(\bar{x},\bar{u};\eta,v) \geq 0$.

$(b) \Rightarrow (c)$: Let (x,u) be any *feasible* pair. It follows that $(\eta,v) := (x-\bar{x}, u-\bar{u})$ is an *admissible* direction. Using condition (b) and the admissiblity of (η,v) in

$$\mathbf{F}(x,u) = \mathbf{F}(\bar{x}+\eta,\bar{u}+v) = \mathbf{F}(\bar{x},\bar{u}) + \mathbf{F}'(\bar{x},\bar{u};\eta,v) + \frac{\mathbf{F}''(\bar{x},\bar{u};\eta,v)}{2},$$

we get $\mathbf{F}(x,u) \geq \mathbf{F}(\bar{x},\bar{u})$, and hence $(\bar{x},\bar{u})$ is a global minimum.

$(c) \Rightarrow (a)$: This implication is immediate. ∎

In order to characterize the primal form of the first-order condition, namely, $\mathbf{F}'(\bar{x},\bar{u};\eta,v) = 0$ for all admissible directions, in terms of the weak maximum principle we introduce the notions of *extremals* and the *weak-normality* for the problem $(\mathcal{LQ}^\sigma)$.

Definition 2.5 (Extremal for $(\mathcal{LQ}^\sigma)$)**.** A feasible pair $(\bar{x},\bar{u})$ is said to be *extremal* for $(\mathcal{LQ}^\sigma)$ if there exist λ_0, $\bar{\lambda} : [a,\rho(b)]_{\mathbb{T}} \to \mathbb{R}^k$, $\bar{\lambda} \in \mathrm{C}_{\mathrm{prd}}$, $\bar{p} : [a,b]_{\mathbb{T}} \to \mathbb{R}^n$, $\bar{p} \in \mathrm{C}^1_{\mathrm{prd}}$, and $\bar{\gamma} \in \mathbb{R}^r$ satisfying

(i) $\lambda_0 \geq 0$, and $\lambda_0 + \|\bar{\lambda}\|_{\mathrm{C}_{\mathrm{prd}}} + \|\bar{p}\|_{\mathrm{C}} + \left|\bar{\gamma}\right| \neq 0$,

(ii) the adjoint equation: for all $t \in [a,\rho(b)]_{\mathbb{T}}$

$$\bar{p}^{\Delta}(t) = -\mathbf{A}^T(t)\bar{p}(t) + \mathbf{K}^T(t)\bar{\lambda}(t) + \lambda_0[\mathbf{P}(t)\bar{x}^\sigma(t) + \mathbf{Q}(t)\bar{u}(t) + \mathbf{z}(t)], \tag{2.10}$$

(iii) the stationarity condition: for all $t \in [a,\rho(b)]_{\mathbb{T}}$

$$-\mathbf{B}^T(t)\,\bar{p}(t) + \mathbf{N}^T(t)\,\bar{\lambda}(t) + \lambda_0[\mathbf{Q}^T(t)\bar{x}^\sigma(t) + \mathbf{R}(t)\bar{u}(t) + \mathbf{k}(t)] = 0, \tag{2.11}$$

(iv) the transversality condition:

$$\begin{pmatrix} \bar{p}(a) \\ -\bar{p}(b) \end{pmatrix} = \lambda_0\Big[\mathbf{\Gamma}\begin{pmatrix} \bar{x}(a) \\ \bar{x}(b) \end{pmatrix} + \begin{pmatrix} \mathbf{c_a} \\ \mathbf{c_b} \end{pmatrix}\Big] + \mathbf{M}^T\bar{\gamma}. \tag{2.12}$$

When $\lambda_0 \neq 0$, it can be taken to be equal to 1.

Remark 2.6. The notion of extremality of a feasible pair means that there exists a vector of multipliers $\bar{\Lambda} := (\lambda_0, \bar{\lambda}, \bar{p}, \bar{\gamma})$ satisfying "the weak maximum principle," that is, conditions (i)–(iv) of Definition 2.5. Conditions (ii)–(iii) can be phrased in terms of the Hamiltonian $\mathbf{H}$ defined in by (1.5) as follows: for all $t \in [a,\rho(b)]_{\mathbb{T}}$,

$$((\bar{p}^{\Delta})^T(t), 0) = \nabla_{x,u}\mathbf{H}(t, \bar{x}^\sigma(t), \bar{u}(t), -\bar{p}(t), \lambda(t), \lambda_0). \tag{2.13}$$

Now we define the weak-normality of our problem at an extremal. It extends the notion known for *continuous-time* control problems with only *pure* control constraints (see e.g., [35, Definition 2.3]).

Definition 2.7 (Weak-Normality of $(\mathcal{LQ}^\sigma)$)**.** The problem $(\mathcal{LQ}^\sigma)$ is *weakly normal* at an extremal pair $(\bar{x}, \bar{u})$ if there exists a multiplier vector of the form $\bar{\Lambda} = (\lambda_0 = 1, \bar{\lambda}, \bar{p}, \bar{\gamma})$ satisfying conditions (i)–(iv) of Definition 2.5.

For comparison purposes only, we introduce here the notions of *normality* and **M**-*controllability*. The normality notion for the problem $(\mathcal{LQ}^\sigma)$ extends to the time scale setting the notion known in [34] for the continuous-time.

Definition 2.8 (Normality of $(\mathcal{LQ}^\sigma)$)**.** The problem $(\mathcal{LQ}^\sigma)$ is *normal* if the system

$$p^\Delta = -\mathbf{A}^T p + \mathbf{K}^T \lambda, -\mathbf{B}^T p + \mathbf{N}^T \lambda = 0, (\text{on } [a, \rho(b)]_{\mathbb{T}}), \begin{pmatrix} p(a) \\ -p(b) \end{pmatrix} = \mathbf{M}^T \gamma, \tag{2.14}$$

where $\gamma \in \mathbb{R}^r$ and $\lambda : [a, \rho(b)]_{\mathbb{T}} \to \mathbb{R}^k$, possesses only the trivial solution $p(\cdot) \equiv 0$ (and then also $\gamma = 0$ and $\lambda \equiv 0$).

Remark 2.9. Similarly to the case of continuous time setting (see [34, Proposition 4.1]), it is easy to see that the normality notion in Definition 2.8 is equivalent to saying that

$$p^\Delta = -(\mathbf{A}^T - \mathbf{K}^T \mathbf{N}_\dagger^T \mathbf{B}^T) p, \mathbf{Y}^T \mathbf{B}^T p = 0, (\text{ on } [a, \rho(b)]_{\mathbb{T}}), \begin{pmatrix} p(a) \\ -p(b) \end{pmatrix} = \mathbf{M}^T \gamma, \tag{2.15}$$

where $\gamma \in \mathbb{R}^r$, possesses only the trivial solution $p(\cdot) \equiv 0$ (and then also $\gamma = 0$).

Next we define the notion of **M**-*controllability* of the system (2.1)–(2.2), which generalizes to the case of *mixed* control-state constraints the notion known, e.g., in [23] where only *pure* control equality constraints are present.

Definition 2.10. The system (2.1)–(2.2) is said to be **M**-*controllable* if

$$\left\{ \mathbf{M} \begin{pmatrix} \eta(a) \\ \eta(b) \end{pmatrix} : (\eta, v) \in \mathrm{C}^1_{\mathrm{prd}} \times \mathrm{C}_{\mathrm{prd}} \text{ and solve } (2.1)\text{–}(2.2) \right\} = \mathbb{R}^r. \tag{2.16}$$

It is equivalent to the **M**-controllability of the control equation (2.6), that is,

$$\left\{ \mathbf{M} \begin{pmatrix} \eta(a) \\ \eta(b) \end{pmatrix} : (\eta, w) \in \mathrm{C}^1_{\mathrm{prd}} \times \mathrm{C}_{\mathrm{prd}} \text{ and solve } (2.6) \right\} = \mathbb{R}^r. \tag{2.17}$$

Remark 2.11 (Normality of $(\mathcal{LQ}^\sigma) \Leftrightarrow$ **M**-controllability)**.** When $I - \mu(\mathbf{A} - \mathbf{B}\mathbf{N}_\dagger \mathbf{K})$ is invertible, arguments similar to [34, Proposition 4.2] and to [23, Proposition 4.5] easily prove that the normality of $(\mathcal{LQ}^\sigma)$ is equivalent to the **M**-controllability of the system (2.1)–(2.2).

When only *pure* control constraints are present, i.e., when $\mathbf{K} \equiv 0$, the above equivalence would only require $I - \mu\mathbf{A}$ to be invertible; which is the usual condition needed for this special setting (see [23, Proposition 4.5]).

Remark 2.12 (Normality vs. weak-normality of $(\mathcal{LQ}^\sigma)$)**.** The normality condition is clearly stronger than that of the weak-normality at an extremal pair $(\bar{x}, \bar{u})$. The former dictates that all multiplier vectors corresponding to any extremal pair of $(\mathcal{LQ}^\sigma)$ must have $\lambda_0 \neq 0$. On the other hand, the weak normality at an extremal $(\bar{x}, \bar{u})$ guarantees for this extremal the existence of at least one multiplier vector having $\lambda_0 \neq 0$. However, this latter notion says nothing about other extremals and does not exclude that there could be for $(\bar{x}, \bar{u})$ other multiplier vectors in which $\lambda_0 = 0$. Note that in many situations, e.g., deriving sufficient optimality criteria, it is enough to have one multiplier vector Λ in which $\lambda_0 \neq 0$.

The following statement confirms that *without* the $\mathbf{M}$-controllability assumption on the system (2.1)–(2.2), the first variation of $(\mathcal{LQ}^\sigma)$ at a pair $(\bar{x}, \bar{u})$ is zero along all *admissible* directions exactly means that $(\bar{x}, \bar{u})$ must be an *extremal* at which the problem $(\mathcal{LQ}^\sigma)$ is *weakly normal*.

Theorem 2.13 (First variation of $(\mathcal{LQ}^\sigma)$ & weak-normality at extremals)**.** *Assume that for all t in $[a, \rho(b)]$ the matrix $I - \mu(\mathbf{A} - \mathbf{B}\mathbf{N}_\dagger\mathbf{K})(t)$ is invertible, and let $(\bar{x}, \bar{u})$ be a feasible pair. Then, the following conditions are equivalent.*

(I) $\mathbf{F}'(\bar{x}, \bar{u}; \eta, v) = 0$ *for all admissible directions* (η, v)*.*
(II) $(\bar{x}, \bar{u})$ *is an extremal at which* $(\mathcal{LQ}^\sigma)$ *is weakly-normal.*

Remark 2.14. When the time scale is *continuous*, $(\mathcal{LQ}^\sigma)$ is *purely* linear-quadratic, both state-endpoints are fixed to be zero (i.e, $\mathbf{M} =$ identity), and the control variable is unrestricted (i.e., (1.3) is absent), then the implication (I) $\Rightarrow$ (II) is proved in [13, Section 8] using Hilbert space methods with the control space being the larger space $L^2[a, b]$. Therefore, the necessity part of Theorem 2.13 generalizes the corresponding result in [13] not only to the time scale setting, but also in the continuous time, to the case where the linear terms of $\mathbf{F}$ are present and where the control-state space is that of piecewise smooth $\times$ piecewise continuous functions, the state-endpoints vary in an affine fashion, and the affine mixed state-control constraints (1.3) are present.

A direct consequence of combining Theorem 2.13 with Proposition 2.3 is the following Corollary, which represents a characterization of optimality of a feasible pair in $(\mathcal{LQ}^\sigma)$. In particular, *without* assuming any controllability condition, an optimal pair $(\bar{x}, \bar{u})$ turns out to be an *extremal* with corresponding multipliers vector having $\lambda_0 \neq 1$. This type of results is a special feature for the linear-quadratic problem, since in the nonlinear problem this feature does not hold without an extra condition like the $\mathbf{M}$-controllability (see [23]).

Corollary 2.15 (Weak-normality of $(\mathcal{LQ}^\sigma)$ at optimal pairs)**.** *Assume that* $I-\mu(\mathbf{A}-\mathbf{B}\mathbf{N}_\dagger\mathbf{K})(t)$ *is invertible on* $[a,\rho(b)]$. *Then, the following conditions are equivalent.*

(a) *The pair* $(\bar{x},\bar{u})$ *is (weak local or global) optimal for* $(\mathcal{LQ}^\sigma)$.

(b) *The pair* $(\bar{x},\bar{u})$ *is an extremal at which the problem* $(\mathcal{LQ}^\sigma)$ *is weakly normal, and* $\mathbf{F}''(\bar{x},\bar{u};\eta,v)\geq 0$ *for all admissible directions* (η,v).

Remark 2.16. When $(\bar{x},\bar{u})$ is an extremal at which $(\mathcal{LQ}^\sigma)$ is weakly normal, that is, the first part of condition (b) holds, with corresponding multipliers $\Lambda=(1,\bar{\lambda},\bar{p},\bar{\gamma})$, then

$$2\mathbf{F}(\bar{x},\bar{u})=\begin{pmatrix}\bar{x}(a)\\ \bar{x}(b)\end{pmatrix}^T\begin{pmatrix}\mathbf{c_a}\\ \mathbf{c_b}\end{pmatrix}-\mathbf{d}^T\bar{\gamma}+\int_a^b[\mathbf{z}^T\bar{x}^\sigma+\mathbf{k}^T\bar{u}-\mathbf{c}^T(t)\bar{\lambda}(t)]\Delta t, \tag{2.18}$$

and hence, if in addition $(\mathcal{LQ}^\sigma)$ is *purely* linear-quadratic, then $\mathbf{F}(\bar{x},\bar{u})=0$.

Proof of Theorem 2.13. (I) $\Rightarrow$ (II): Assume (I), i.e., $\mathbf{F}'(\bar{x},\bar{u};\eta,v)=0$ for all admissible directions (η,v). As discussed at the beginning of this section, (η,v) is admissible means that v is given by (2.5) for some $w\in \mathrm{C_{prd}}$, and (η,w) satisfies (2.6)–(2.7).

Let $\Phi(\cdot)$ to be the fundamental matrix of the system (2.6), i.e., $\mathbf{\Phi}(t)$ be the unique solution of $Z^\Delta=\Big(\mathbf{A}(t)-\mathbf{B}(t)\mathbf{N}_\dagger(t)\mathbf{K}(t)\Big)Z^\sigma$, $Z(a)=I$. By [6, Theorem 5.21] and the invertibility of $I-\mu\Big(\mathbf{A}-\mathbf{B}\mathbf{N}_\dagger\mathbf{K}\Big)(t)$ on $[a,\rho(b)]_{\mathbb{T}}$, the matrix $\mathbf{\Phi}(t)$ exists and is invertible on $[a,b]_{\mathbb{T}}$. The fundamental matrix $\mathbf{\Psi}$ of the adjoint system

$$Z^\Delta=-\Big(\mathbf{A}(t)-\mathbf{B}(t)\mathbf{N}_\dagger(t)\mathbf{K}(t)\Big)^TZ,\quad Z(a)=I, \tag{2.19}$$

satisfies $\mathbf{\Phi}^T(t)\mathbf{\Psi}(t)=I$ on $[a,b]_{\mathbb{T}}$.

Set $\mathbf{M}=[\mathbf{M}_a\mathbf{M}_b]$, where $\mathbf{M}_a$ and $\mathbf{M}_b$ are respectively the first and the last n-columns of $\mathbf{M}$. By using the variation of constant formula, see e.g. [6, Theorem 5.27], where we write $\eta(a)=\alpha$, it follows that the solutions of (2.6) and (2.7) are of the form

$$\eta(t)=\mathbf{\Phi}(t)\alpha+\mathbf{\Phi}(t)\int_a^t\mathbf{\Phi}^{-1}(\tau)\mathbf{B}(\tau)\mathbf{Y}(\tau)w(\tau)\Delta\tau, \tag{2.20}$$

$$[\mathbf{M}_a+\mathbf{M}_b\mathbf{\Phi}(b)]\alpha+\mathbf{M}_b\mathbf{\Phi}(b)\int_a^b\mathbf{\Phi}^{-1}(\tau)\mathbf{B}(\tau)\mathbf{Y}(\tau)w(\tau)\Delta\tau=0, \tag{2.21}$$

where $\alpha\in\mathbb{R}^n$ and $w\in\mathrm{C_{prd}}$.

Write $\mathbf{\Gamma}=\begin{bmatrix}\mathbf{\Gamma}_a & \mathbf{\Gamma}_{ab}\\ \mathbf{\Gamma}_{ba} & \mathbf{\Gamma}_b\end{bmatrix}$, where each entry is an $n\times n$-matrix and substitute $\eta(a)=\alpha$, $v=\mathbf{Y}w-\mathbf{N}_\dagger\mathbf{K}\eta^\sigma$, and η^σ from (2.20), into the equation $\mathbf{F}'(\bar{x},\bar{u};\eta,v)=0$. It results that, for

$$\beta_1:=\mathbf{\Gamma}_a\bar{x}(a)+\mathbf{\Gamma}_{ba}\bar{x}(b)+\mathbf{c_a},\qquad \beta_2:=\mathbf{\Gamma}_{ab}\bar{x}(a)+\mathbf{\Gamma}_b\bar{x}(b)+\mathbf{c_b}, \tag{2.22}$$

$$\left.\begin{aligned}\delta(t) &:= \Big[(\mathbf{P}-\mathbf{K}^T\mathbf{N}_\dagger^T\mathbf{Q}^T)\bar{x}^\sigma + (\mathbf{Q}-\mathbf{K}^T\mathbf{N}_\dagger^T\mathbf{R})\bar{u} + \mathbf{z} - \mathbf{K}^T\mathbf{N}_\dagger^T\mathbf{k}\Big](t),\\ \xi(t) &:= \mathbf{Q}(t)^T\bar{x}^\sigma(t) + \mathbf{R}(t)\bar{u}(t) + \mathbf{k}(t),\end{aligned}\right\} \tag{2.23}$$

we have

$$\begin{aligned}0 = {}& [\beta_1^T + \beta_2^T\mathbf{\Phi}(b) + \int_a^b \delta^T(t)\mathbf{\Phi}^\sigma(t)\Delta t]\alpha\\ &+ \int_a^b [\beta_2^T\mathbf{\Phi}(b)\mathbf{\Phi}^{-1}(t)\mathbf{B}(t) + \xi^T(t)]\mathbf{Y}(t)w(t)\Delta t\\ &+ \int_a^b \delta^T(t)\mathbf{\Phi}^\sigma(t)\Big(\int_a^{\sigma(t)} \mathbf{\Phi}^{-1}(\tau)\mathbf{B}(\tau)\mathbf{Y}(\tau)w(\tau)\Delta\tau\Big)\Delta t,\end{aligned} \tag{2.24}$$

for all $\alpha \in \mathbb{R}^n$ and $w \in \mathrm{C}_{\mathrm{prd}}$ satisfying (2.21).

To calculate the last term of (2.24) we use the integration by parts formula

$$\int_a^b y^\Delta(t)z^\sigma(t)\Delta t = y(b)z(b) - y(a)z(a) - \int_a^b y(t)z^\Delta(t)\Delta t,$$

where $y(t) := \int_a^t \delta^T(\tau)\mathbf{\Phi}^\sigma(\tau)\Delta\tau$, and $z(t) := \int_a^t \mathbf{\Phi}^{-1}(\tau)\mathbf{B}(\tau)\mathbf{Y}(\tau)w(\tau)\Delta\tau$ we get

$$\begin{aligned}&\int_a^b \delta^T(t)\mathbf{\Phi}^\sigma(t)\Big(\int_a^{\sigma(t)} \mathbf{\Phi}^{-1}(\tau)\mathbf{B}(\tau)\mathbf{Y}(\tau)w(\tau)\Delta\tau\Big)\Delta t\\ &\qquad = \int_a^b \delta^T(\tau)\mathbf{\Phi}^\sigma(\tau)\Delta\tau \int_a^b \mathbf{\Phi}^{-1}(t)\mathbf{B}(t)\mathbf{Y}(t)w(t)\Delta t\\ &\qquad\quad - \int_a^b \Big(\int_a^t \delta^T(\tau)\mathbf{\Phi}^\sigma(\tau)\Delta\tau\Big)\mathbf{\Phi}^{-1}(t)\mathbf{B}(t)\mathbf{Y}(t)w(t)\Delta t\\ &\qquad = \int_a^b \Big(\int_t^b \delta^T(\tau)\mathbf{\Phi}^\sigma(\tau)\Delta\tau\Big)\mathbf{\Phi}^{-1}(t)\mathbf{B}(t)\mathbf{Y}(t)w(t)\Delta t\end{aligned} \tag{2.25}$$

Using (2.25) into (2.24) we obtain

$$\begin{aligned}0 = {}& [\beta_1^T + \beta_2^T\mathbf{\Phi}(b) + \int_a^b \delta^T(t)\mathbf{\Phi}^\sigma(t)\Delta t]\alpha\\ &+ \int_a^b \Big[\Big(\beta_2^T\mathbf{\Phi}(b) + \int_t^b \delta^T(\tau)\mathbf{\Phi}^\sigma(\tau)\Delta\tau\Big)\mathbf{\Phi}^{-1}(t)\mathbf{B}(t) + \xi^T(t)\Big]\mathbf{Y}(t)w(t)\Delta t,\end{aligned} \tag{2.26}$$

for all $\alpha \in \mathbb{R}^n$ and $w : [a, \rho(b)] \to \mathbb{R}^{m-k}$ satisfying (2.21). Set

$$\left\{\begin{aligned}&\mathcal{D} := \mathbf{M}_a + \mathbf{M}_b\mathbf{\Phi}(b) \in \mathbb{R}^{r\times n},\\ &\mathcal{E}(t) := \mathbf{M}_b\mathbf{\Phi}(b)\mathbf{\Phi}^{-1}(t)\mathbf{B}(t)\mathbf{Y}(t) \in \mathbb{R}^{r\times(m-k)},\\ &d := \beta_1 + \mathbf{\Phi}^T(b)\beta_2 + \textstyle\int_a^b(\mathbf{\Phi}^\sigma)^T(t)\delta(t)\Delta t \in \mathbb{R}^n,\\ &h(t) := \mathbf{Y}^T(t)\Big[\mathbf{B}^T(t)\mathbf{\Phi}^{-T}(t)\Big(\mathbf{\Phi}^T(b)\beta_2\\ &\qquad + \textstyle\int_t^b(\mathbf{\Phi}^\sigma)^T(\tau)\delta(\tau)\Delta\tau\Big) + \xi(t)\Big] \in \mathbb{R}^{m-k}.\end{aligned}\right. \tag{2.27}$$

Then, (2.26) and (2.21) are equivalent to saying that

$$d^T\alpha + \int_a^b h^T(t)w(t)\Delta t = 0, \quad \text{whenever} \quad \mathcal{D}\alpha + \int_a^b \mathcal{E}(t)w(t)\Delta t = 0, \tag{2.28}$$

for $\alpha \in \mathbb{R}^n$ and $w : [a, \rho(b)]_{\mathbb{T}} \to \mathbb{R}^{m-k}$, $w \in \mathrm{C}_{\mathrm{prd}}$. Thus, by the *generalized Dubois-Reymond Lemma* in [23, Lemma 4.7], (2.28) is equivalent to the existence of a vector $c \in \mathbb{R}^r$ such that

$$d = \mathcal{D}^T c \quad \text{ and } \quad h(t) = \mathcal{E}^T(t) c \quad \text{for all } t \in [a, \rho(b)]_{\mathbb{T}}.$$

This means that

$$\beta_1 + \mathbf{\Phi}^T(b)\beta_2 + \int_a^b (\mathbf{\Phi}^\sigma)^T(t)\delta(t)\Delta t = [\mathbf{M}_a^T + \mathbf{\Phi}^T(b)\mathbf{M}_b^T]c, \tag{2.29}$$

$$\begin{aligned} &\mathbf{Y}^T(t)\Big[\mathbf{B}^T(t)\mathbf{\Phi}^{-T}(t)\Big(\mathbf{\Phi}^T(b)\beta_2 + \int_t^b (\mathbf{\Phi}^\sigma)^T(\tau)\delta(\tau)\Delta\tau\Big) + \xi(t)\Big] \\ &\quad = \mathbf{Y}^T(t)\mathbf{B}^T(t)\mathbf{\Phi}^{-T}(t)\mathbf{\Phi}^T(b)\mathbf{M}_b^T c, \qquad \forall\, t \in [a, \rho(b)]_{\mathbb{T}}. \end{aligned} \tag{2.30}$$

Set $\bar{\gamma} := -c$ and

$$\bar{p}(t) := \mathbf{\Phi}^{-T}(t)\Big[-\mathbf{\Phi}^T(b)\beta_2 - \int_t^b (\mathbf{\Phi}^\sigma)^T(\tau)\delta(\tau)\Delta\tau + \mathbf{\Phi}^T(b)\mathbf{M}_b^T c\Big]. \tag{2.31}$$

From (2.31), (2.29), and (2.22), it easily follows that $\bar{p}$ and $\bar{\gamma}$ satisfy the transversality conditions (2.12) where $\lambda_0 = 1$.

Now, (2.31) and (2.23) yield that (2.30) is equivalent to

$$\mathbf{Y}^T(t)[\mathbf{B}^T(t)\bar{p}(t) - \mathbf{Q}^T(t)\bar{x}^\sigma(t) - \mathbf{R}(t)\bar{u}(t) - \mathbf{k}(t)] = 0, \quad \forall\, t \in [a, \rho(b)], \tag{2.32}$$

which is equivalent to the existence of a function $\bar{\lambda} : [a, \rho(b)] \to \mathbb{R}^k$ such that

$$\mathbf{B}^T(t)\bar{p}(t) - \mathbf{Q}^T(t)\bar{x}^\sigma(t) - \mathbf{R}(t)\bar{u}(t) - \mathbf{k}(t) = \mathbf{N}^T(t)\bar{\lambda}(t), \quad \forall\, t \in [a, \rho(b)], \tag{2.33}$$

proving that (2.11) is satisfied by $\lambda_0 = 1$, $\bar{p}$ and $\bar{\lambda}$. Furthermore, the full rank property of $\mathbf{N}$ implies the uniqueness of the function $\bar{\lambda}$ satisfying (2.33) and that it must be given by the formula

$$\bar{\lambda}(t) = \mathbf{N}_\dagger^T(t)[\mathbf{B}^T(t)\bar{p}(t) - \mathbf{Q}^T(t)\bar{x}^\sigma(t) - \mathbf{R}(t)\bar{u}(t) - \mathbf{k}(t)] \quad \forall\, t \in [a, \rho(b)], \tag{2.34}$$

where $\mathbf{N}_\dagger$ is defined in (2.4). This proves that $\bar{\lambda}$ is indeed in $\mathrm{C}_{\mathrm{prd}}$.

It remains to show that $\lambda_0 = 1$, $\bar{p}$, and $\bar{\lambda}$ satisfy the adjoint equation. Calculate the Δ-derivative of $\bar{p}$ given by (2.31), and using (2.23) for $\delta(t)$ and the fact that $\mathbf{\Psi} = \mathbf{\Phi}^{-T}$ satisfies (2.19) we obtain

$$\bar{p}^\Delta = -\mathbf{A}^T\bar{p} + \mathbf{P}\bar{x}^\sigma + \mathbf{Q}\bar{u} + \mathbf{z} + \mathbf{K}^T\mathbf{N}_\dagger^T\Big[\mathbf{B}^T\bar{p} - \mathbf{Q}^T\bar{x}^\sigma - \mathbf{R}\bar{u} - \mathbf{k}\Big], \quad \text{on } [a, \rho(b)]. \tag{2.35}$$

Using the definition of $\bar{\lambda}$ in (2.34) into (2.35) we conclude that the adjoint equation (2.10) holds. Therefore, Condition (II) of the theorem is satisfied.

(II) $\Rightarrow$ (I): Assume that there is a multiplier vector $\Lambda := (\lambda_0 = 1, \bar{\lambda}, \bar{p}, \bar{\gamma})$ satisfying with the pair $(\bar{x}, \bar{u})$ conditions (ii)–(iv) of (2.5). Let (η, v) be any admissible pair, from (2.10)–(2.12), and (2.1)–(2.3) it follows that the expression of $\mathbf{F}'$ in $(\mathcal{L}^\sigma)$ is

$$\begin{aligned}
\mathbf{F}'(\bar{x},\bar{u};\eta,v) :&= {\mathbf{c_a}}^T\eta(a) + {\mathbf{c_b}}^T\eta(b) + \begin{pmatrix}\bar{x}(a)\\ \bar{x}(b)\end{pmatrix}^T \boldsymbol{\Gamma} \begin{pmatrix}\eta(a)\\ \eta(b)\end{pmatrix} \\
&\quad + \int_a^b \Big[\Big((\bar{x}^\sigma)^{!T}\mathbf{P} + \bar{u}^T\mathbf{Q}^T + \mathbf{z}^T\Big)\eta^\sigma + \Big((\bar{x}^\sigma)^T\mathbf{Q} + \bar{u}^T\mathbf{R} + \mathbf{k}^T\Big)v\Big]\Delta t \\
&= \Big[\begin{pmatrix}\bar{p}(a)\\ -\bar{p}(b)\end{pmatrix}^T - \bar{\gamma}^T\mathbf{M}\Big]\begin{pmatrix}\eta(a)\\ \eta(b)\end{pmatrix} \\
&\quad + \int_a^b \Big[(\bar{p}^\Delta)^T\eta^\sigma + \bar{p}^T(\mathbf{A}\eta^\sigma + \mathbf{B}v) - \bar{\lambda}^T(\mathbf{K}\eta^\sigma + \mathbf{N}v)\Big]\Delta t \\
&= 0.
\end{aligned}$$

$\blacksquare$

Remark 2.17. Note that in the above proof, the invertibility assumption on $I - \mu(\mathbf{A} - \mathbf{B}\mathbf{N}_\dagger\mathbf{K})$ is only needed to prove the implication $(I) \Rightarrow$ (II). However, this assumption is not needed to prove (II) $\Rightarrow$ (I).

In this last part of this section, we will show that under the invertibility of a certain matrix function $\mathbf{S}$, the *weak normality* of $(\mathcal{LQ}^\sigma)$ at an extremal $(\bar{x},\bar{u})$ with adjoint variable $\bar{p}$ is equivalent to $(\bar{x},\bar{p})$ solving a *nonhomogeneous symplectic boundary value problem* and to $\bar{u}$ being a certain matrix linear combination of $(\bar{x},\bar{p})$. Thereby, we will be able via Corollary 2.15 to characterize the optimality in terms of the existence of a solution to this symplectic boundary value problem.

Let $\mathbf{Y} : [a,\rho(b)]_{\mathbb{T}} \to \mathbb{R}^{m\times(m-k)}$ be the matrix function introduced in the proof of Theorem 2.13. Thus, $(\bar{x},\bar{u})$ feasible, that is, satisfying (1.2)–(1.4), is equivalent to saying that for some function $\bar{w} : [a,\rho(b)]_{\mathbb{T}} \to \mathbb{R}^{(m-k)}$, $\bar{w} \in C_{prd}$

$$\bar{x}^\Delta = (\mathbf{A} - \mathbf{B}\mathbf{N}_\dagger\mathbf{K})\bar{x}^\sigma + \mathbf{B}(\mathbf{Y}\bar{w} + \mathbf{N}_\dagger\mathbf{c}), \tag{2.36}$$

$$\bar{u} = \mathbf{Y}\bar{w} + \mathbf{N}_\dagger(\mathbf{c} - \mathbf{K}\bar{x}^\sigma), \tag{2.37}$$

$$\mathbf{M}\begin{pmatrix}\bar{x}(a)\\ \bar{x}(b)\end{pmatrix} = \mathbf{d}. \tag{2.38}$$

Assume that for all t, $(I - \mu(\mathbf{A} - \mathbf{B}\mathbf{N}_\dagger\mathbf{K}))(t)$ is *invertible*, and define $\tilde{\mathbf{C}}(t)$ to be its inverse, that is,

$$\tilde{\mathbf{C}}(t) := (I - \mu(\mathbf{A} - \mathbf{B}\mathbf{N}_\dagger\mathbf{K}))^{-1}(t). \tag{2.39}$$

In this case (2.36) and (2.37) are equivalent to

$$\bar{x}^\sigma = \tilde{\mathbf{C}}\bar{x} + \mu\tilde{\mathbf{C}}\mathbf{B}(\mathbf{Y}\bar{w} + \mathbf{N}_\dagger\mathbf{c}), \tag{2.40}$$

$$\bar{x}^\Delta = (\mathbf{A} - \mathbf{B}\mathbf{N}_\dagger\mathbf{K})\tilde{\mathbf{C}}\bar{x} + \tilde{\mathbf{C}}\mathbf{B}(\mathbf{Y}\bar{w} + \mathbf{N}_\dagger\mathbf{c}), \tag{2.41}$$

$$\bar{u} = (I - \mu\mathbf{N}_\dagger\mathbf{K}\tilde{\mathbf{C}}\mathbf{B})(\mathbf{Y}\bar{w} + \mathbf{N}_\dagger\mathbf{c}) - \mathbf{N}_\dagger\mathbf{K}\tilde{\mathbf{C}}\bar{x}. \tag{2.42}$$

Define the $(m-k)\times(m-k)$-matrix function $\mathbf{S}$ on $[a,\rho(b)]_{\mathbb{T}}$ by

$$\mathbf{S}(t) := \mathbf{Y}^T(t)\Big[\mathbf{R} + \mu(\mathbf{Q}^T - \mathbf{R}\mathbf{N}_\dagger\mathbf{K})\tilde{\mathbf{C}}\mathbf{B}\Big](t)\mathbf{Y}(t). \tag{2.43}$$

When $\mathbf{S}(t)$ is invertible for all t, define the *non-homogeneous symplectic boundary value problem* on $[a,\rho(b)]_{\mathbb{T}}$ by

$$(\mathbb{S}) \qquad \begin{cases} \bar{x}^\Delta = \mathbb{A}(t)\bar{x} + \mathbb{B}(t)\bar{p} + \tilde{\mathbf{C}}\mathbf{B}\zeta(t) \\ \bar{p}^\Delta = \mathbb{C}(t)\bar{x} + \mathbb{D}(t)\bar{p} + \rho(t) \end{cases}$$

$$(\text{Bdry}) \qquad \mathbf{M}\begin{pmatrix}\bar{x}(a)\\ \bar{x}(b)\end{pmatrix} = \mathbf{d}, \quad \begin{pmatrix}\bar{p}(a)\\ -\bar{p}(b)\end{pmatrix} = \mathbf{\Gamma}\begin{pmatrix}\bar{x}(a)\\ \bar{x}(b)\end{pmatrix} + \begin{pmatrix}\mathbf{c_a}\\ \mathbf{c_b}\end{pmatrix} + \mathbf{M^T}\,\bar{\gamma},$$

where

$$\begin{cases} \zeta := \mathbf{N}_\dagger\mathbf{c} - \mathbf{Y}\mathbf{S}^{-1}\mathbf{Y}^T\Big[\mathbf{k} + \Big(\mathbf{R} + \mu(\mathbf{Q}^T - \mathbf{R}\mathbf{N}_\dagger\mathbf{K})\tilde{\mathbf{C}}\mathbf{B}\Big)\mathbf{N}_\dagger\mathbf{c}\Big], \\ \rho := \mathbf{z} - \mathbf{K}^T\mathbf{N}_\dagger^T\mathbf{k} + [\mathbf{Q} + \mu(\mathbf{P} - \mathbf{Q}\mathbf{N}_\dagger\mathbf{K})\tilde{\mathbf{C}}\mathbf{B} - \mathbf{K}^T\mathbf{N}_\dagger^T(\mathbf{R} + \mu(\mathbf{Q}^T - \mathbf{R}\mathbf{N}_\dagger\mathbf{K})\tilde{\mathbf{C}}\mathbf{B}]\zeta, \\ \mathbb{A} := [(\mathbf{A} - \mathbf{B}\mathbf{N}_\dagger\mathbf{K} - \tilde{\mathbf{C}}\mathbf{B}\mathbf{Y}\mathbf{S}^{-1}\mathbf{Y}^T(\mathbf{Q}^T - \mathbf{R}\mathbf{N}_\dagger\mathbf{K})]\tilde{\mathbf{C}}, \\ \mathbb{B} := \tilde{\mathbf{C}}\mathbf{B}\mathbf{Y}\mathbf{S}^{-1}\mathbf{Y}^T\mathbf{B}^T, \\ \mathbb{C} := \Big[(\mathbf{P} - (\mathbf{Q} - \mathbf{K}^T\mathbf{N}_\dagger^T\mathbf{R})\mathbf{N}_\dagger\mathbf{K} - \mathbf{K}^T\mathbf{N}_\dagger^T\mathbf{Q}^T)(I - \mu\tilde{\mathbf{C}}\mathbf{B}\mathbf{Y}\mathbf{S}^{-1}\mathbf{Y}^T(\mathbf{Q}^T - \mathbf{R}\mathbf{N}_\dagger\mathbf{K})) \\ \qquad + (\mathbf{K}^T\mathbf{N}_\dagger^T\mathbf{R} - \mathbf{Q})\mathbf{Y}\mathbf{S}^{-1}\mathbf{Y}^T(\mathbf{Q}^T - \mathbf{R}\mathbf{N}_\dagger\mathbf{K})\Big]\tilde{\mathbf{C}}, \\ \mathbb{D} := \Big(\mathbf{P} - \mathbf{Q}\mathbf{N}_\dagger\mathbf{K} - \mathbf{K}^T\mathbf{N}_\dagger^T(\mathbf{Q}^T - \mathbf{R}\mathbf{N}_\dagger\mathbf{K})\Big)\mu\tilde{\mathbf{C}}\mathbf{B}\mathbf{Y}\mathbf{S}^{-1}\mathbf{Y}^T\mathbf{B}^T \\ \qquad + (\mathbf{Q} - \mathbf{K}^T\mathbf{N}_\dagger^T\mathbf{R})\mathbf{Y}\mathbf{S}^{-1}\mathbf{Y}^T\mathbf{B}^T + \mathbf{K}^T\mathbf{N}_\dagger^T\mathbf{B}^T - \mathbf{A}^T. \end{cases} \tag{2.44}$$

In the continuous time setting (i.e., $\mu \equiv 0$), and when the initial endpoint is fixed, system $(\mathbb{S})$ and (Bdry) reduce to to system (5.8)–(5.11) in [34].

When the mixed state-control constraints is absent (i.e., $\mathbf{K} \equiv 0$ and $\mathbf{N} \equiv 0$), no endpoint constraints are present (i.e., $\mathbf{M} = 0$ and $\mathbf{d} = 0$), and $\mathbf{F}$ is *purely* quadratic, the coefficients (2.44) and system $(\mathbb{S})$ reduce to the coefficients and their corresponding *homogeneous* symplectic system obtained in [24, Theorem 3.1]. Similarly to the proof of [24, Theorem 3.1], the coefficients in (2.44) actually define a time scale *symplectic system* since it can be verified that they satisfy

$$\begin{cases} \mathbb{C}^T(I + \mu\mathbb{A}) \text{ and } (I + \mu\mathbb{D}^T)\mathbb{B} \text{ are symmetric,} \\ \text{and } \mathbb{A}^T + \mathbb{D} + \mu(\mathbb{A}^T\mathbb{D} - \mathbb{C}^T\mathbb{B}) = 0, \end{cases} \tag{2.45}$$

and

$$\begin{cases} \mathbb{B}(I + \mu\mathbb{A})^T \text{ and } \mathbb{C}(I + \mu\mathbb{D}^T) \text{ are symmetric,} \\ \text{and } \mathbb{A}^T + \mathbb{D} + \mu(\mathbb{D}\mathbb{A}^T - \mathbb{C}\mathbb{B}^T) = 0. \end{cases} \tag{2.46}$$

Proposition 2.18 (Weak-normality and symplectic systems for $(\mathcal{LQ}^\sigma)$)**.** *Assume that* $[I-\mu(\mathbf{A}-\mathbf{B}\mathbf{N}_\dagger\mathbf{K})](t)$ *and* $\mathbf{S}(t)$ *are invertible for all* t *in* $[a,\rho(b)]$*. Then, the following conditions are equivalent.*

(a) *The problem* $(\mathcal{LQ}^\sigma)$ *is weakly-normal at the extremal* $(\bar{x},\bar{u})$ *with associated* $\bar{\Lambda} := (1,\bar{\lambda},\bar{p},\bar{\gamma})$.

(b) $(\bar{x},\bar{p},\bar{\gamma})$ *solve the symplectic boundary value problem* $(\mathbb{S})$ *and (Bdry), and*

$$\begin{aligned}\bar{u} = (I-\mu\mathbf{N}_\dagger\mathbf{K}\tilde{\mathbf{C}}\mathbf{B})\mathbf{Y}\mathbf{S}^{-1}\mathbf{Y}^T[\mathbf{B}^T\bar{p}-(\mathbf{Q}^T-\mathbf{R}\mathbf{N}_\dagger\mathbf{K})\tilde{\mathbf{C}}\bar{x}] \\ -\mathbf{N}_\dagger\mathbf{K}\tilde{\mathbf{C}}\bar{x}+(I-\mu\mathbf{N}_\dagger\mathbf{K}\tilde{\mathbf{C}}\mathbf{B})\zeta,\end{aligned} \tag{2.47}$$

where ζ *is defined by* (2.44).

Remark 2.19. When $(\mathcal{LQ}^\sigma)$ is *purely* linear-quadratic, then $\zeta\equiv 0$, $\rho\equiv 0$, and $\mathbf{c}_a=\mathbf{c}_b=0$. In this case, the corresponding symplectic system $(\mathbb{S})$ is *homogeneous* and the boundary conditions (Bdry) take the form

$$\mathbf{M}\begin{pmatrix}\bar{x}(a)\\ \bar{x}(b)\end{pmatrix}=0,\qquad \begin{pmatrix}\bar{p}(a)\\ -\bar{p}(b)\end{pmatrix}=\mathbf{\Gamma}\begin{pmatrix}\bar{x}(a)\\ \bar{x}(b)\end{pmatrix}+\mathbf{M}^T\bar{\gamma}. \tag{2.48}$$

Proof of Proposition 2.18*.* Assume (a) to be valid. Then, (2.10)–(2.12) are satisfied for $\lambda_0=1$, and by the feasibility of $(\bar{x},\bar{u})$, (2.40)–(2.41) and (2.38) hold for some $\bar{w}:[a,\rho(b)]_{\mathbb{T}}\to\mathbb{R}^{(m-k)}$, $\bar{w}\in C_{prd}$. Substitute for $\bar{x}^\sigma$ and $\bar{u}$ from (2.40) and (2.42) into (2.10) and (2.11), where $\lambda_0=1$, we obtain

$$\bar{p}^\Delta=-\mathbf{A}^T\bar{p}+\mathbf{K}^T\bar{\lambda}+(\mathbf{P}-\mathbf{Q}\mathbf{N}_\dagger\mathbf{K})\tilde{\mathbf{C}}\bar{x}+\Big[\mathbf{Q}+\mu(\mathbf{P}-\mathbf{Q}\mathbf{N}_\dagger\mathbf{K})\tilde{\mathbf{C}}\mathbf{B}\Big](\mathbf{Y}\bar{w}+\mathbf{N}_\dagger\mathbf{c})+\mathbf{z}, \tag{2.49}$$

$$-\mathbf{B}^T\bar{p}+\mathbf{N}^T\bar{\lambda}+(\mathbf{Q}^T-\mathbf{R}\mathbf{N}_\dagger\mathbf{K})\tilde{\mathbf{C}}\bar{x}+\Big[\mathbf{R}+\mu(\mathbf{Q}^T-\mathbf{R}\mathbf{N}_\dagger\mathbf{K})\tilde{\mathbf{C}}\mathbf{B}\Big](\mathbf{Y}\bar{w}+\mathbf{N}_\dagger\mathbf{c})+\mathbf{k}=0. \tag{2.50}$$

By the full rank property of $\mathbf{N}$, and from the definition of $\mathbf{S}$ in (2.43), it follows that Equation (2.50) is equivalent to

$$\mathbf{Y}^T(-\mathbf{B}^T\bar{p}+\mathbf{k})+\mathbf{Y}^T(\mathbf{Q}^T-\mathbf{R}\mathbf{N}_\dagger\mathbf{K})\tilde{\mathbf{C}}\bar{x}+\mathbf{S}\bar{w}+\mathbf{Y}^T[\mathbf{R}+\mu(\mathbf{Q}^T-\mathbf{R}\mathbf{N}_\dagger\mathbf{K})\tilde{\mathbf{C}}\mathbf{B}]\mathbf{N}_\dagger\mathbf{c}=0, \tag{2.51}$$

$$\bar{\lambda}=\mathbf{N}_\dagger^T\Big[\mathbf{B}^T\bar{p}-\mathbf{k}-(\mathbf{Q}^T-\mathbf{R}\mathbf{N}_\dagger\mathbf{K})\tilde{\mathbf{C}}\bar{x}-\Big(\mathbf{R}+\mu(\mathbf{Q}^T-\mathbf{R}\mathbf{N}_\dagger\mathbf{K})\tilde{\mathbf{C}}\mathbf{B}\Big)(\mathbf{Y}\bar{w}+\mathbf{N}_\dagger\mathbf{c})\Big], \tag{2.52}$$

where $\mathbf{N}_\dagger$ is defined in (2.4). The invertibility of $\mathbf{S}(t)$, for all t, yields that (2.51) and (2.52) allow the functions $\bar{w}$ and $\bar{\lambda}$ to be phrased in terms of $(\bar{x},\bar{p})$ as follows:

$$\bar{w}=\mathbf{S}^{-1}\mathbf{Y}^T\Big[\mathbf{B}^T\bar{p}-\mathbf{k}-(\mathbf{Q}^T-\mathbf{R}\mathbf{N}_\dagger\mathbf{K})\tilde{\mathbf{C}}\bar{x}-\mathscr{S}\mathbf{N}_\dagger\mathbf{c}\Big], \tag{2.53}$$

$$\bar{\lambda}=\mathbf{N}_\dagger^T\Big[I-\mathscr{S}\mathbf{Y}\mathbf{S}^{-1}\mathbf{Y}^T\Big]\Big[\mathbf{B}^T\bar{p}-\mathbf{k}-(\mathbf{Q}^T-\mathbf{R}\mathbf{N}_\dagger\mathbf{K})\tilde{\mathbf{C}}\bar{x}-\mathscr{S}\mathbf{N}_\dagger\mathbf{c}\Big], \tag{2.54}$$

where

$$\mathscr{S} := \mathbf{R} + \mu(\mathbf{Q}^T - \mathbf{R}\mathbf{N}_\dagger \mathbf{K})\tilde{\mathbf{C}}\mathbf{B}. \tag{2.55}$$

Now, substitute (2.53) and (2.54) into (2.42), (2.41), and (2.49), we obtain that $\bar{u}$ satisfies (2.47) and $(\bar{x}, \bar{p}, \bar{\gamma})$ satisfy the *symplectic system* $(\mathbb{S})$ and the boundary conditions (Bdry).

Conversely, assume (b) is true, that is, $(\bar{x}, \bar{p}, \bar{\gamma})$ solve $(\mathbb{S})$ and (Bdry) and $\bar{u}$ is given by (2.47). Define $\bar{w}$ and $\bar{\lambda}$ respectively by (2.53) and (2.52). Then, it follows that $(\bar{x}, \bar{p}, \bar{w}, \bar{\gamma})$ satisfy (2.41), (2.49), (2.50), (2.38), and (Bdry). This means that (a) is satisfied. ∎

Remark 2.20. Substituting the formula (2.37) for $\bar{u}$ in terms of $\bar{w}$, $\bar{p}$ and $\bar{x}^\sigma$ into the stationarity equation (2.11) with $\lambda_0 = 1$, we obtain an equation in $\bar{w}$, $\bar{p}$, and $\bar{x}^\sigma$, namely,

$$-\mathbf{B}^T\bar{p} + \mathbf{N}^T\bar{\lambda} + (\mathbf{Q}^T - \mathbf{R}\mathbf{N}_\dagger\mathbf{K})\bar{x}^\sigma + \mathbf{R}\mathbf{N}_\dagger\mathbf{c} + \mathbf{k} + \mathbf{R}\mathbf{Y}\bar{w} = 0. \tag{2.56}$$

In order to solve this last equation for $\bar{w}$ in terms of $(\bar{x}^\sigma, \bar{p})$ we would need to assume as in [25] that $\mathbf{Y}^T\mathbf{R}\mathbf{Y}$ is invertible. However, this is too restrictive, since in the discrete case this assumption is rarely met. For this reason, in the proof above, the stationarity equation was phrased in terms of $\bar{x}$ (instead of $\bar{x}^\sigma$) in which $\mathbf{S}$ turns out to be the coefficient of $\bar{w}$ (see (2.50)). While in the continuous time setting $\mathbf{S}$ is $\mathbf{R}$, this is not the case in the discrete time setting. Therefore, in the above proposition, the invertibility of only $\mathbf{S}$, and not of $\mathbf{R}$, is of great importance for the case when $\mu(t) \neq 0$. It allowed us to go directly from the *Jacobi system* $(\mathbf{J}^\sigma)$, defined by (1.2)–(1.4) and (2.10)–(2.12) (where $\lambda_0 = 1$), to the Symplectic System $(\mathbb{S})$ and (Bdry) without having to go through the "Hamiltonian System" (see [24] for the case where no mixed state-control or endpoints constraints are present and $(\mathcal{LQ}^\sigma)$ is *purely* linear-quadratic).

Now by using Proposition 2.18 we are able to re-state Corollary 2.15 in terms of the symplectic system $(\mathbb{S})$.

Corollary 2.21 (Characterization of optimality and symplectic system for $(\mathcal{LQ}^\sigma)$)**.** *Assume* $[I - \mu(\mathbf{A} - \mathbf{B}\mathbf{N}_\dagger\mathbf{K})](t)$ *and* $\mathbf{S}(t)$ *are invertible for all* t *in* $[a, \rho(b)]$. *Then, the following conditions are equivalent.*

- (a) $(\bar{x}, \bar{u})$ *is a (weak local or global) optimal for* $(\mathcal{LQ}^\sigma)$.
- (b) *There exist a function* $\bar{p}$ *and a vector* $\bar{\gamma}$ *such that* $(\bar{x}, \bar{p}, \bar{\gamma})$ *solve* $(\mathbb{S})$ *and* (Bdry), $\bar{u}$ *is given via* (2.47), *and* $\mathbf{F}''(\bar{x}, \bar{u}; \eta, v) \geq 0$ *for all admissible pairs.*

Remark 2.22. Assume that $\mathbf{F}$ is *purely* linear-quadratic, then the set of admissible directions (η, v) coincides with that of feasible pairs and $2\mathbf{F}(\eta, v) = \mathbf{F}''(\bar{x}, \bar{u}; \eta, v)$.

In the continuous time setting, the non-negativity and positivity of $\mathbf{F}''(\bar{x},\bar{u};\eta,v)$ are characterized by conditions in terms of the notion of *focal points* and the *Riccati equation* corresponding to the Jacobi system (see e.g, [31]–[35]). In the discrete time setting (see e.g., [14]–[16], [18] and [26]) and in the time scale setting (see e.g., [17], [19], [20]), the question of characterizing the nonnegativity and positivity of $\mathbf{F}''(\bar{x},\bar{u};\eta,v)$ was intensively studied when $\mathbf{F}$ itself is *symplectic* (i.e., of the form ($\mathbf{D}$) below), or when we are in the calculus of variations setting. However, this question remains open for quadratic functionals of the form ($\mathcal{LQ}^\sigma$) (with or without control-state constraints).

To the symplectic system ($\mathbb{S}$) and its boundary conditions (Bdry) obtained from ($\mathcal{LQ}^\sigma$) we can associate a quadratic *diagonal functional* $\mathbb{F}(x,p)$ and a *diagonal* linear-quadratic problem ($\mathbf{D}$) defined as:

$$\text{minimize} \quad \mathbb{F}(x,p) := \mathbf{c_a}^T x(a) + \mathbf{c_b}^T x(b) + \frac{1}{2}\begin{pmatrix} x(a)\\ x(b)\end{pmatrix}^T \mathbf{\Gamma} \begin{pmatrix} x(a)\\ x(b)\end{pmatrix} + \int_a^b \Big[\underline{\mathbf{z}}^T x + \underline{\mathbf{k}}^T p$$

$$(\mathbf{D}) \qquad + \frac{1}{2}\Big(x^T\mathbb{C}^T(I+\mu\mathbb{A})x + 2\mu x^T\mathbb{C}^T\mathbb{B}p + p^T(I+\mu\mathbb{D}^T)\mathbb{B}p\Big)\Big](t)\Delta t$$

subject to: $x \in \mathrm{C}^1_{\mathrm{prd}}[a,b]_{\mathbb{T}}$ and $u \in \mathrm{C}_{\mathrm{prd}}[a,\rho(b)]_{\mathbb{T}}$ satisfying

$$x^\Delta(t) = \mathbb{A}(t)x(t) + \mathbb{B}(t)p(t) + \tilde{\mathbf{C}}\mathbf{B}\zeta, \quad t \in [a,\rho(b)]_{\mathbb{T}}, \tag{2.57}$$

$$\mathbf{M}\begin{pmatrix} x(a)\\ x(b)\end{pmatrix} = \mathbf{d}, \tag{2.58}$$

where the coefficient ζ, $\mathbb{A},\mathbb{B},\mathbb{C}$, and $\mathbb{D}$ are defined through (2.44), and hence satisfy (2.45) and (2.46), and where $\underline{\mathbf{z}}$ and $\underline{\mathbf{k}}$ are

(2.59)
$$\underline{\mathbf{z}} := (I+\mu\mathbb{A}^T)\mathbf{z} - \tilde{\mathbf{C}}^T\Big[\mathbf{K^T N_\dagger^T} + (\mathbf{Q} - \mathbf{K^T N_\dagger^T R})\mathbf{Y S^{-T} Y^T}(\mathbf{I} - \mu\mathbf{B^T}\tilde{\mathbf{C}}^\mathbf{T}\mathbf{K^T N_\dagger^T})\Big]\mathbf{k},$$

$$\underline{\mathbf{k}} := \mu\mathbb{B}^T\mathbf{z} + \mathbf{B Y S^{-T} Y^T}(\mathbf{I} - \mu\mathbf{B^T}\tilde{\mathbf{C}}^\mathbf{T}\mathbf{K^T N_\dagger^T})\mathbf{k}. \tag{2.60}$$

Remark 2.23. Since ($\mathbf{D}$) is a linear-quadratic problem of *symplectic* structure, one can show (see Section 4) that the symplectic system corresponding to its second variation $\mathbb{F}''$ is exactly ($\mathbb{S}$) itself, in which $\zeta \equiv 0$ and $\rho \equiv 0$, and its boundary conditions (Bdry), where $\mathbf{c}_a = \mathbf{c}_b = 0$. Given that there are extensive results in the literature characterizing the non-negativity and positivity of $\mathbb{F}''$, one would wonder whether those results could be transferred to the original problem ($\mathcal{LQ}^\sigma$) and its second variation $\mathbf{F}''$. This motivates an important question: are the problems ($\mathcal{LQ}^\sigma$) and ($\mathbf{D}$) equivalent? If so, then checking the nonnegativity (or positivity) for $\mathbf{F}''(\bar{x},\bar{u};\eta,v)$ over the *admissible directions* (η,v), would be equivalent to that of $\mathbb{F}''(\bar{x},\bar{p};\xi,q)$ over (ξ,q) satisfying (2.57) and (2.58), where $\zeta \equiv 0$, and $\mathbf{d} = 0$. By Corollary 2.21, it follows that even in the general case one implication is true, namely, if $(\bar{x},\bar{u})$ is optimal for $\mathbf{F}(x,u)$

so does $(\bar{x}, \bar{p})$ for $\mathbb{F}(x, p)$, where $\bar{p}$ is the function defined through its part (b). In fact, if (x, p) is feasible for $(\mathbf{D})$ then, define u in terms of (x, p) by (2.47), we then have the pair (x, u) feasible for $(\mathcal{LQ}^\sigma)$ and that $\mathbf{F}(x, u) = \mathbb{F}(x, p) + [\mathbf{k}^T + \mu(\mathbf{z}^T - \mathbf{k}^T \mathbf{N}_\dagger \mathbf{K})\tilde{\mathbf{C}}\mathbf{B}]\zeta$. Similarly, we have $\mathbf{F}''(\bar{x}, \bar{u}; \eta, v) = \mathbb{F}''(\bar{x}, \bar{p}; \xi, q)$. However, the converse does not appear to be true without extra assumption. If fact, for a given pair (x, u) feasible for $(\mathcal{LQ}^\sigma)$ we would need to get p with (x, p) feasible for $(\mathbf{D})$. Solving (2.56) or its equivalent equation (2.50) for p in terms of (x, u) would require the assumption that for all t, $(\mathbf{Y}^T\mathbf{B})(t)$ is of full rank; which is too strong and is not commonly satisfied in control problems.

3. LINEAR-QUADRATIC PROBLEMS WITHOUT STATE SHIFT

In this section we derive for the linear-quadratic control problem $(\mathcal{LQ})$ introduced in Section 1 results parallel to those derived in Section 2 for the problem $(\mathcal{LQ}^\sigma)$.

Definition 3.1. A pair (η, v) is said to be an *admissible direction* for the problem $(\mathcal{LQ})$ if it satisfies the equation of motion (1.7), and the linearized equations of (1.8) and (1.9), that is,

$$\eta^\Delta(t) = A(t)\eta(t) + B(t)v(t), \quad t \in [a, \rho(b)]_{\mathbb{T}}, \tag{3.1}$$

$$K(t)\eta(t) + N(t)v(t) = 0, \quad t \in [a, \rho(b)]_{\mathbb{T}}, \tag{3.2}$$

$$M \begin{pmatrix} \eta(a) \\ \eta(b) \end{pmatrix} = 0. \tag{3.3}$$

Note that if the $c \equiv 0$ and $d = 0$ then feasible pairs and admissible directions are the same.

Since $N(t)$ is of full rank for all $t \in [a, \rho(b)]_{\mathbb{T}}$, choose a function $Y : [a, \rho(b)]_{\mathbb{T}} \to \mathbb{R}^{m \times (m-k)}$, $Y \in \mathrm{C}_{\mathrm{prd}}$, such that, for each t, $Y(t)$ is a matrix whose columns form an orthonormal basis for the space $\operatorname{Ker} N(t)$. Set

$$N_\dagger(t) := N^T(t)(N(t)N^T(t))^{-1} \quad \text{for all } t \in [a, \rho(b)]_{\mathbb{T}}. \tag{3.4}$$

Then, equation (3.2) is equivalent to

$$v(t) = Y(t)w(t) - N_\dagger(t)K(t)\eta(t), \quad \text{for all } t \in [a, \rho(b)]_{\mathbb{T}}, \tag{3.5}$$

where $w(\cdot)$ is in $\mathrm{C}_{\mathrm{prd}}$. Hence, (η, v) satisfies Definition (3.1) is equivalent to (η, w) satisfies

$$\eta^\Delta(t) = \Big(A(t) - B(t)N_\dagger(t)K(t)\Big)\eta(t) + B(t)Y(t)w(t), \quad t \in [a, \rho(b)]_{\mathbb{T}}, \tag{3.6}$$

$$M \begin{pmatrix} \eta(a) \\ \eta(b) \end{pmatrix} = 0, \tag{3.7}$$

where v is defined through w in (3.5).

We define the first and second variations of the problem $(\mathcal{LQ})$ at pair $(\bar{x}, \bar{u})$ in $\mathrm{C}^1_{\mathrm{prd}} \times \mathrm{C}_{\mathrm{prd}}$ in the direction of an admissible pair (η, v) by

$$F'(\bar{x}, \bar{u}; \eta, v) := {c_a}^T \eta(a) + {c_b}^T \eta(b) + \begin{pmatrix} \bar{x}(a) \\ \bar{x}(b) \end{pmatrix}^T \Gamma \begin{pmatrix} \eta(a) \\ \eta(b) \end{pmatrix}$$

$$(\mathcal{L}) \qquad + \int_a^b \Big[z^T \eta + k^T v + \bar{x}^T P \eta + \bar{u}^T Q^T \eta + \bar{x}^T Q\, v + \bar{u}^T R v \Big](t)\, \Delta t,$$

$$(\mathcal{Q}) \qquad F''(\bar{x}, \bar{u}; \eta, v) := \begin{pmatrix} \eta(a) \\ \eta(b) \end{pmatrix}^T \Gamma \begin{pmatrix} \eta(a) \\ \eta(b) \end{pmatrix} + \int_a^b [\eta^T P \eta + 2\eta^T Q v + v^T R v](t) \Delta t.$$

The following proposition states that, the necessary conditions of first and second order do not require neither the invertibility of $I + \mu(A - BN_\dagger K)$ nor the M-controllability of (3.1)–(3.2). In addition, this Proposition shows that these necessary conditions are also sufficient for the weak local optimality of $(\bar{x}, \bar{u})$, which is equivalent to the global optimality.

Proposition 3.2 (Optimality & 1st and 2nd variations of $(\mathcal{LQ})$)**.** *Let $(\bar{x}, \bar{u})$ be feasible for $(\mathcal{LQ})$. Then, the following conditions are equivalent.*

(a) $(\bar{x}, \bar{u})$ *is a weak local minimum for* $(\mathcal{LQ})$.
(b) *For all admissible directions* (η, v), $F'(\bar{x}, \bar{u}; \eta, v) = 0$ *and* $F''(\bar{x}, \bar{u}; \eta, v) \geq 0$.
(c) $(\bar{x}, \bar{u})$ *is a global minimum for* $(\mathcal{LQ})$.

Proof. The proof of this proposition is identical to that of Proposition 2.3 where $\mathbf{F}, \mathbf{F}'$ and $\mathbf{F}''$ are respectively replaced by F, F' and F''. ∎

Parallel to Definitions 2.5 and 2.7, we introduce here the notions of extremals and *weak normality* for problem $(\mathcal{LQ})$ in terms of which the first-order condition is characterized.

Definition 3.3 (Extremal for $(\mathcal{LQ})$)**.** A feasible pair $(\bar{x}, \bar{u})$ is said to be *extremal* for $(\mathcal{LQ})$ if there exist λ_0, $\bar{\lambda} : [a, \rho(b)]_{\mathbb{T}} \to \mathbb{R}^k$, $\bar{\lambda} \in \mathrm{C}_{\mathrm{prd}}$, $\bar{p} : [a, b]_{\mathbb{T}} \to \mathbb{R}^n$, $\bar{p} \in \mathrm{C}^1_{\mathrm{prd}}$, and $\bar{\gamma} \in \mathbb{R}^r$ satisfying

(i) $\lambda_0 \geq 0$, and $\lambda_0 + \|\bar{\lambda}\|_{\mathrm{C}_{\mathrm{prd}}} + \|\bar{p}\|_{\mathrm{C}} + |\bar{\gamma}| \neq 0$,
(ii) the adjoint equation: for all $t \in [a, \rho(b)]_{\mathbb{T}}$

$$(3.8) \qquad \bar{p}^\Delta(t) = -A^T(t) \bar{p}^\sigma(t) + K^T(t) \bar{\lambda}(t) + \lambda_0 [P(t) \bar{x}(t) + Q(t) \bar{u}(t) + z(t)],$$

(iii) the stationarity condition: for all $t \in [a, \rho(b)]_{\mathbb{T}}$

$$(3.9) \qquad -B^T(t) \bar{p}^\sigma(t) + N^T(t) \bar{\lambda}(t) + \lambda_0 [Q^T(t) \bar{x}(t) + R(t) \bar{u}(t) + k(t)] = 0,$$

(iv) the transversality condition:

$$\begin{pmatrix} \bar{p}(a) \\ -\bar{p}(b) \end{pmatrix} = \lambda_0 \Big[\Gamma \begin{pmatrix} \bar{x}(a) \\ \bar{x}(b) \end{pmatrix} + \begin{pmatrix} c_a \\ c_b \end{pmatrix} \Big] + M^T \bar{\gamma}. \tag{3.10}$$

Definition 3.4 (Weak-Normality of $(\mathcal{LQ})$)**.** The problem $(\mathcal{LQ})$ is *weakly normal* at an extremal pair $(\bar{x}, \bar{u})$ if there exists a multiplier vector of the form $\bar{\Lambda} = (\lambda_0 = 1, \bar{\lambda}, \bar{p}, \bar{\gamma})$ satisfying conditions (i)–(iv) of Definition 3.3.

For comparison purposes only, we introduce here the notions of *normality* and *M-controllability* in $(\mathcal{LQ})$.

Definition 3.5 (Normality of $(\mathcal{LQ})$)**.** The problem $(\mathcal{LQ})$ is *normal* if the system

$$p^{\Delta} = -A^T p^{\sigma} + K^T \lambda, -B^T p + N^T \lambda = 0, (\text{on } [a, \rho(b)]_{\mathbb{T}}), \begin{pmatrix} p(a) \\ -p(b) \end{pmatrix} = M^T \gamma, \tag{3.11}$$

where $\gamma \in \mathbb{R}^r$ and $\lambda : [a, \rho(b)]_{\mathbb{T}} \to \mathbb{R}^k$, possesses only the trivial solution $p(\cdot) \equiv 0$ (and then also $\gamma = 0$ and $\lambda \equiv 0$).

Remark 3.6. Similarly to the case of continuous time setting (see [34, Proposition 4.1]), it is easy to see that the normality notion in Definition (2.8) is equivalent to

$$p^{\Delta} = -(A^T - K^T N_{\dagger}^T B^T)p, Y^T B^T p = 0, (\text{on } [a, \rho(b)]_{\mathbb{T}}), \begin{pmatrix} p(a) \\ -p(b) \end{pmatrix} = M^T \gamma, \tag{3.12}$$

where $\gamma \in \mathbb{R}^r$, possesses only the trivial solution $p(\cdot) \equiv 0$ (and then also $\gamma = 0$).

Definition 3.7. The system (3.1)-(3.2) is said to be *M-controllable* if

$$\Big\{ M \begin{pmatrix} \eta(a) \\ \eta(b) \end{pmatrix} : (\eta, v) \in \mathrm{C}^1_{\mathrm{prd}} \times \mathrm{C}_{\mathrm{prd}} \text{ and solve (3.1)–(3.2)} \Big\} = \mathbb{R}^r. \tag{3.13}$$

It is equivalent to the M-controllability of the control equation (3.6), that is,

$$\Big\{ M \begin{pmatrix} \eta(a) \\ \eta(b) \end{pmatrix} : (\eta, w) \in \mathrm{C}^1_{\mathrm{prd}} \times \mathrm{C}_{\mathrm{prd}} \text{ and solve (3.6)} \Big\} = \mathbb{R}^r. \tag{3.14}$$

Remark 3.8 (Normality of $(\mathcal{LQ}) \Leftrightarrow M$-controllability)**.** When $I + \mu(A - B\, N_{\dagger}\, K)$ is invertible, arguments parallel to [34, Proposition 4.2] and to [23, Theorem 9.3], easily prove that the normality of $(\mathcal{LQ})$ is equivalent to the M-controllability of the system (3.1)–(3.2).

Remark 3.9 (Normality vs. weak-normality in $(\mathcal{LQ})$)**.** It is clear that the notion of *normality* for problem $(\mathcal{LQ})$ is stronger than that of the *weak normality* at an extremal pair $(\bar{x}, \bar{u})$. In many situations including sufficient optimality criteria we only require to have one multiplier in which $\lambda_0 \neq 0$, and hence it suffices to have weak normality.

Parallel to Theorem 2.13, the following result confirms that *without* any assumption of M-controllability on the linear system, the first variation at a pair $(\bar{x}, \bar{u})$ of $(\mathcal{LQ})$ along all *admissible* directions is zero, means that $(\bar{x}, \bar{u})$ must be an *extremal* at which the problem $(\mathcal{LQ})$ is *weakly normal.*

Theorem 3.10 (First variation of $(\mathcal{LQ})$ & weak-normalilty at extremals)**.** *Assume that $I+\mu(A-BN_\dagger K)$ is invertible on $[a,\rho(b)]$, and let $(\bar{x}, \bar{u})$ be a feasible pair. Then, the following conditions are equivalent.*

(i) *$F'(\bar{x}, \bar{u}; \eta, v) = 0$ for all admissible pairs (η, v).*
(ii) *$(\bar{x}, \bar{u})$ is an extremal at which $(\mathcal{LQ})$ is weakly normal.*

Prior to proving Theorem 3.10, we state the following Corollary, obtained by combining Proposition 3.2 with Theorem 3.10. It is a characterization for the optimality in $(\mathcal{LQ})$ in terms of the weak- normality of the problem and the non-negativity of the second variation.

Corollary 3.11 (Weak-normality of $(\mathcal{LQ})$ at optimal pairs)**.** *Assume that $I+\mu(A-BN_\dagger K)$ is invertible on $[a,\rho(b)]$. Then, the following conditions are equivalent.*

(a) *$(\bar{x}, \bar{u})$ is optimal (weak local or global) for $(\mathcal{LQ})$.*
(b) *The pair $(\bar{x}, \bar{u})$ is an extremal at which the problem $(\mathcal{LQ})$ is weakly-normal, and $F''(\bar{x}, \bar{u}; \eta, v) \geq 0$ for all admissible directions (η, v).*

Remark 3.12. Corollary 3.11 implies that when we merely assume that $I + \mu(A - B\,N_\dagger\,K)$ is invertible on $[a, \rho(b)]$, the linear-quadratic problem $(\mathcal{LQ})$ is *weakly-normal* at optimal solutions. Unlike [23, Theorem 9.4], which applies when $K \equiv 0$, no M-controllability is needed to obtain at least one multiplier in which $\lambda_0 \neq 0$.

The proof of Theorem 3.10 could be done directly by employing arguments parallel to those used in the proof of Theorem 2.13. However, a more constructive proof of the result is presented here in two steps: (1) by establishing the equivalence between $(\mathcal{LQ})$ and$(\mathcal{LQ}^\sigma)$, and then (2) by writing problem $(\mathcal{LQ})$ in the form of $(\mathcal{LQ}^\sigma)$ to which Theorem 2.13 applies and whose results are translated back in terms of the data of $(\mathcal{LQ})$.

The following two propositions show that the two forms of the linear quadratic problem are equivalent. These statements are instrumental when we know a result for one form and we wish to prove parallel result for the other form.

Proposition 3.13 ($(\mathcal{LQ})$ to $(\mathcal{LQ}^\sigma)$)**.** *Let P, Q, R, A, B, K, N, c, z, k, Γ, M, c_a, c_b, and d be the data defining Problem $(\mathcal{LQ})$ and let $N_\dagger$ be defined by* (3.4)*. Assume that $I + \mu(A - B\,N_\dagger\,K)$ is invertible on $[a, \rho(b)]$ with inverse $\tilde{C}$. Set the associated* *problem* $(\mathcal{LQ}^\sigma)$ *with coefficients*

$$\left\{\begin{aligned} &\mathbf{\Gamma} := \Gamma; \quad \mathbf{M} := M; \quad \mathbf{c_a} := c_a; \quad \mathbf{c_b} := c_b; \quad \mathbf{d} := d; \\ &\mathbf{A} := \tilde{C}\,A; \quad \mathbf{B} := \tilde{C}\,B; \quad \mathbf{K} := K; \quad \mathbf{N} := N; \quad \mathbf{c} := c; \\ &\begin{pmatrix} \mathbf{P} & \mathbf{Q} \\ \mathbf{Q}^T & \mathbf{R} \end{pmatrix} := \Omega^T \begin{pmatrix} P & Q \\ Q^T & R \end{pmatrix} \Omega; \quad \begin{pmatrix} \mathbf{z} \\ \mathbf{k} \end{pmatrix} = \Omega^T \begin{pmatrix} z \\ k \end{pmatrix}, \end{aligned}\right. \tag{3.15}$$

where

$$\Omega := \begin{pmatrix} I - \mu\tilde{C}A & -\mu\tilde{C}B \\ \mu N_\dagger K\tilde{C}A & I + \mu N_\dagger K\tilde{C}B \end{pmatrix} \ \textit{and}\ \Omega^{-1} = \begin{pmatrix} I + \mu A & \mu B \\ -\mu N_\dagger K A & I - \mu N_\dagger K B \end{pmatrix} \tag{3.16}$$

Then,

(a) $\tilde{\mathbf{C}}^{-1} := I - \mu(\mathbf{A} - \mathbf{B}\mathbf{N}_\dagger\mathbf{K}) = \tilde{C}$;

(b) (x, u) *feasible (respectively,* (η, v) *admissible) for* $(\mathcal{LQ})$ *if and only if* (x, ω) *is feasible (respectively,* (η, ν) *is admissible) for* $(\mathcal{LQ}^\sigma)$*, where*

$$\omega := u - \mu N_\dagger K x^\Delta, \quad \textit{and} \quad \nu := v - \mu N_\dagger K \eta^\Delta. \tag{3.17}$$

Furthermore,

$$\left\{\begin{aligned} &\begin{pmatrix} x \\ u \end{pmatrix} = \Omega \begin{pmatrix} x^\sigma \\ \omega \end{pmatrix}; \quad \begin{pmatrix} \eta \\ v \end{pmatrix} = \Omega \begin{pmatrix} \eta^\sigma \\ \nu \end{pmatrix}; \\ &F(x, u) = \mathbf{F}(x, \omega); \quad F'(x, u; \eta, v) = \mathbf{F}'(x, w; \eta, \nu). \end{aligned}\right. \tag{3.18}$$

Remark 3.14. For the case where only pure control constraints are present, that is, $K \equiv 0$, and when the function F is *purely* quadratic, this result follows from [23, Proposition 3.1]. However, taking $K \equiv 0$ simplifies the results a great deal, since in this case the control ω defined in (3.17) would be exactly u. That is, if $K \equiv 0$ and (x, u) is feasible in $(\mathcal{LQ})$ then, under $I + \mu A$ invertible, the equation of motion (1.7) can easily be phrased in terms of x^σ and u as

$$x^\Delta = \tilde{A}Ax^\sigma + \tilde{A}Bu, \tag{3.19}$$

where $\tilde{A} = (I + \mu A)^{-1}$. However, when K is present, using (3.19) in the control-state constraint (1.8) yields

$$K\tilde{A}x^\sigma + (N - \mu K\tilde{A}B)u = c.$$

This equation has $\mathbf{N} := N - \mu K\tilde{A}B$, which is not necessarily of full rank. Thus, we would not be able to obtain the result of Theorem 3.10 by applying Theorem 2.13 to the quadratic problem with shift in the state obtained in this manner. Therefore, the difficulty in this proposition is to obtain a form of the control ω in terms of the pair (x, u) so that the feasibility equations for (x, u) and the functional $F(x, u)$ translate in terms of (x^σ, ω) in such a way that (x, ω) be feasible for the translated problem $(\mathcal{LQ}^\sigma)$ and that the transition matrix Ω be invertible.

Proof of Proposition 3.13. From the definitions of $\mathbf{A}$, $\mathbf{B}$, $\mathbf{N}_\dagger$ and $\mathbf{K}$ in (3.15) and by setting $\tilde{C} := [I + \mu(A - B\, N_\dagger\, K)]^{-1}$, conclusion (a) of the proposition follows.

Let (x, u) be feasible for $(\mathcal{LQ})$ and let $\omega := u - \mu N_\dagger K x^\Delta$. Then, $Nu = N\omega + \mu K x^\Delta$. Substituting into (1.7)–(1.8) the equations

$$u = \omega + \mu N_\dagger K x^\Delta \quad \text{and} \quad x = x^\sigma - \mu x^\Delta, \tag{3.20}$$

we obtain

$$x^\Delta = \tilde{C} A x^\sigma + \tilde{C} B \omega, \quad K x^\sigma + N\omega = c. \tag{3.21}$$

Thus, since x also satisfies (1.9), it results that (x, ω) is feasible for $(\mathcal{LQ}^\sigma)$. Use (3.21)(a) into (3.20) we get, for Ω defined in (3.16), that

$$\begin{pmatrix} x \\ u \end{pmatrix} = \Omega \begin{pmatrix} x^\sigma \\ \omega \end{pmatrix}. \tag{3.22}$$

Conversely, let (x, ω) be a feasible pair for the problem $(\mathcal{LQ}^\sigma)$ whose coefficients are given through (3.15). Then, in addition to (1.9), we have

$$x^\Delta = \tilde{C} A x^\sigma + \tilde{C} B \omega \tag{3.23}$$

$$K x^\sigma + N\omega = c. \tag{3.24}$$

Define $u := \omega + \mu N_\dagger K x^\Delta$, and use $x^\sigma = x + \mu x^\Delta$ in (3.23) it follows that (x, u) is feasible for $(\mathcal{LQ})$ and it satisfies with (x, ω) equation (3.22).

Now substitute (x, u) from (3.22) into $F(x, u)$ defined by (1.6), we readily obtain $F(x, u) = \mathbf{F}(x, \omega)$, where $\mathbf{F}$ is defined by (1.1) with coefficients from (3.15).

Let (η, v) admissible for $(\mathcal{LQ})$. Similarly to the calculation for (x, u) where now $c \equiv 0$ and $d = 0$, we obtain the equivalence with (η, ν) being admissible for $(\mathcal{LQ}^\sigma)$, where ν defined in (3.17)(ii). Moreover, we have

$$\begin{pmatrix} \eta \\ v \end{pmatrix} = \Omega \begin{pmatrix} \eta^\sigma \\ \nu \end{pmatrix}. \tag{3.25}$$

Direct substitution of (3.22) and (3.25) into $F'(x, u; \eta, v)$ defined by $(\mathcal{L})$, we obtain $F'(x, u; \eta, v) = \mathbf{F}'(x, \omega; \eta, \nu)$, where $\mathbf{F}'$ is defined by $(\mathcal{L}^\sigma)$. Hence, Part (2) of this proposition is valid. Furthermore, since $\mathbf{N} := N$ is of full rank, it follows that the Problem $(\mathcal{LQ}^\sigma)$ with the data given by (3.15) is indeed one of the quadratic problems with shift in the state. ■

Parallel to Proposition 3.15, we now start with a Problem of the form $(\mathcal{LQ}^\sigma)$ and translate it to a problem of the form $(\mathcal{LQ})$. The proof of this proposition is similar to the previous one and hence, is omitted.

Proposition 3.15 (($\mathcal{LQ}^\sigma$) to ($\mathcal{LQ}$))**.** *Let* $\mathbf{P}$, $\mathbf{Q}$, $\mathbf{R}$, $\mathbf{A}$, $\mathbf{B}$, $\mathbf{K}$, $\mathbf{N}$, $\mathbf{c}$, $\mathbf{z}$, $\mathbf{k}$, $\mathbf{\Gamma}$, $\mathbf{M}$, $\mathbf{c_a}$, $\mathbf{c_b}$, *and* $\mathbf{d}$ *be the data defining Problem* ($\mathcal{LQ}^\sigma$), *and let* $\mathbf{N}_\dagger$ *be defined by* (2.4)*. Assume that* $I - \mu(\mathbf{A} - \mathbf{B}\,\mathbf{N}_\dagger\,\mathbf{K})$ *is invertible on* $[a, \rho(b)]$ *with inverse* $\tilde{\mathbf{C}}$*. Set the problem* ($\mathcal{LQ}$) *with coefficients*

$$\left\{\begin{array}{l} \Gamma := \mathbf{\Gamma}; \quad M := \mathbf{M}; \quad c_a := \mathbf{c_a}; \quad c_b := \mathbf{c_b}; \quad d := \mathbf{d}; \\ A := \tilde{\mathbf{C}}\mathbf{A}; \quad B := \tilde{\mathbf{C}}\mathbf{B}; \quad K := \mathbf{K}; \quad N := \mathbf{N}; \quad c := \mathbf{c}; \\ \begin{pmatrix} P & Q \\ Q^T & R \end{pmatrix} := \mathbf{\Sigma}^T \begin{pmatrix} \mathbf{P} & \mathbf{Q} \\ \mathbf{Q}^T & \mathbf{R} \end{pmatrix} \mathbf{\Sigma}; \quad \begin{pmatrix} z \\ k \end{pmatrix} = \mathbf{\Sigma}^T \begin{pmatrix} \mathbf{z} \\ \mathbf{k} \end{pmatrix}, \end{array}\right. \tag{3.26}$$

where

$$\mathbf{\Sigma} := \begin{pmatrix} I + \mu\tilde{\mathbf{C}}\mathbf{A} & \mu\tilde{\mathbf{C}}\mathbf{B} \\ -\mu\mathbf{N}_\dagger\mathbf{K}\tilde{\mathbf{C}}\mathbf{A} & I - \mu\mathbf{N}_\dagger\mathbf{K}\tilde{\mathbf{C}}\mathbf{B} \end{pmatrix} \text{ and } \mathbf{\Sigma^{-1}} = \begin{pmatrix} I - \mu\mathbf{A} & -\mu\mathbf{B} \\ \mu\mathbf{N}_\dagger\mathbf{K}\mathbf{A} & I + \mu\mathbf{N}_\dagger\mathbf{K}\mathbf{B} \end{pmatrix} \tag{3.27}$$

Then,

1. $I + \mu(A - BN_\dagger K) = \tilde{\mathbf{C}}$;
2. (x, ω) *feasible (respectively,* (η, ν) *admissible) for* ($\mathcal{LQ}^\sigma$) *if and only if* (x, u) *is feasible (respectively,* (η, v) *is admissible) for* ($\mathcal{LQ}$), *where*

$$u := \omega + \mu\mathbf{N}_\dagger\mathbf{K}x^\Delta \quad \text{and} \quad v := \nu + \mu\mathbf{N}_\dagger\mathbf{K}\eta^\Delta. \tag{3.28}$$

Furthermore,

$$\left\{\begin{array}{l} \begin{pmatrix} x^\sigma \\ \omega \end{pmatrix} = \mathbf{\Sigma}\begin{pmatrix} x \\ u \end{pmatrix}; \quad \begin{pmatrix} \eta^\sigma \\ \nu \end{pmatrix} = \mathbf{\Sigma}\begin{pmatrix} \eta \\ v \end{pmatrix}; \\ \mathbf{F}(x, \omega) = F(x, u); \quad \mathbf{F}'(x, \omega; \eta, \nu) = F'(x, u; \eta, v). \end{array}\right. \tag{3.29}$$

Remark 3.16. In the absence of the mixed state-control constraints (1.3) in ($\mathcal{LQ}^\sigma$) and (1.8) in ($\mathcal{LQ}$), and when $\mathbf{F}$ and F are *purely* quadratic, the coefficients (3.15) and (3.26) reduce respectively to those in Propositions 4.14 and 4.15 in [24].

Proof of Theorem 3.10*.* Define ($\mathcal{LQ}^\sigma$) from Problem ($\mathcal{LQ}$) through (3.15), and set

$$\bar{\omega} := \bar{u} - \mu N_\dagger K \bar{x}^\Delta. \tag{3.30}$$

It results from Proposition 3.13 that $(\bar{x}, \bar{\omega})$ is feasible for ($\mathcal{LQ}^\sigma$) and that

$$\begin{pmatrix} \bar{x} \\ \bar{u} \end{pmatrix} = \Omega \begin{pmatrix} \bar{x}^\sigma \\ \bar{\omega} \end{pmatrix}. \tag{3.31}$$

Part (b) of Proposition 3.13 yields that condition (i) of Theorem 3.10 is equivalent to saying that $\mathbf{F}'(\bar{x}, \bar{\omega}; \eta, \nu) = 0$, for all (η, ν) admissible for the problem ($\mathcal{LQ}^\sigma$) and that for $v = \nu + \mu N_\dagger K \eta^\Delta$, we have (η, v) admissible for ($\mathcal{LQ}$). Thus, Condition (i) of this theorem is equivalent to Condition (I) of Theorem 2.13. Furthermore, from part(a) of Proposition 3.13 we have the invertibility of $\tilde{\mathbf{C}}$, and hence by applying

Theorem 2.13 to our defined problem $(\mathcal{LQ}^\sigma)$, we obtain that Condition (i) of this Theorem is also equivalent to saying that $(\bar{x}, \bar{\omega})$ is an extremal for $(\mathcal{LQ}^\sigma)$ at which the problem $(\mathcal{LQ}^\sigma)$ is weakly-normal. This means that there exist $\bar{\Lambda} := (\lambda_0 = 1, \bar{\lambda}, \bar{p}, \bar{\gamma})$ where $\bar{\lambda} : [a, \rho(b)]_{\mathbb{T}} \to \mathbb{R}^k$, $\bar{\lambda} \in \mathrm{C}_{\mathrm{prd}}$, $\bar{p} : [a, b]_{\mathbb{T}} \to \mathbb{R}^n$, $\bar{p} \in \mathrm{C}^1_{\mathrm{prd}}$, and $\bar{\gamma} \in \mathbb{R}^r$ satisfying (2.10)–(2.12). By means of (3.15) and (3.31), these equations translate in terms of the data of $(\mathcal{LQ})$ to (3.10) with $\lambda_0 = 1$ and, on $[a, \rho(b)]_{\mathbb{T}}$,

$$\bar{p}^\Delta = -A^T \tilde{C}^T \bar{p} + K^T \bar{\lambda} + P\bar{x} + Q\bar{u} + z + \mu A^T \tilde{C}^T \mathcal{W}, \tag{3.32}$$

$$-B^T \tilde{C}^T \bar{p} + N^T \bar{\lambda} + Q^T \bar{x} + R\bar{u} + k + \mu B^T \tilde{C}^T \mathcal{W} = 0, \tag{3.33}$$

where,

$$\mathcal{W} := K^T N_\dagger^T (Q^T \bar{x} + R\bar{u} + k) - (P\bar{x} + Q\bar{u} + z). \tag{3.34}$$

Multiply equation (3.33) by $(-K^T N_\dagger^T)$ and add (3.32) we obtain after using $N_\dagger^T N^T = I$ and $\mu(A^T - K^T N_\dagger^T B^T)\tilde{C}^T = \tilde{C}^T \mu(A^T - K^T N_\dagger^T B^T) = I - \tilde{C}^T$, that

$$\bar{p}^\Delta = -\tilde{C}^T (A^T - K N_\dagger^T B^T)\bar{p} - \tilde{C}^T \mathcal{W}. \tag{3.35}$$

Substitute $\bar{p} = \bar{p}^\sigma - \mu\bar{p}^\Delta$ in (3.35) and solve for $\bar{p}^\Delta$ using that $\tilde{C} = [I + \mu(A - BN_\dagger K)]^{-1}$, it follows that

$$\bar{p}^\Delta = -(A^T - K^T N_\dagger^T B^T)\bar{p}^\sigma - \mathcal{W}. \tag{3.36}$$

Substitute $\bar{p} = \bar{p}^\sigma - \mu\bar{p}^\Delta$ in (3.33) and use (3.36) and the definition of $\tilde{C}$ we get

$$-B^T \bar{p}^\sigma + N^T \bar{\lambda} + Q^T \bar{x} + R\bar{u} + k = 0, \tag{3.37}$$

that is, (3.9) holds with $\lambda_0 = 1$. Now multiply (3.37) by $(K^T N_\dagger^T)$ and use it in (3.36), where $\mathcal{W}$ is given by (3.34), it follows that (3.8) holds with $\lambda_0 = 1$.

The implication (ii) $\Rightarrow$ (i) is straightforward. Therefore, the statement of this theorem is proved. ∎

In this last part of this section, we will show that under the invertibility of a certain matrix function S, the *weak normality* of $(\mathcal{LQ})$ at an extremal $(\bar{x}, \bar{u})$ can be expressed in terms of a non-homogeneous *symplectic boundary value problem*. Thereby, we will be able via Corollary 3.11 to characterize the optimality in terms of the existence of a solution to this symplectic boundary value problem.

Since $c \not\equiv 0$, then, for $(\bar{x}, \bar{u})$ feasible we have parallel to (3.5) and (3.6), that equations (1.7)–(1.9) are equivalent to

$$\bar{x}^\Delta = \Big(A - BN_\dagger K\Big)\bar{x} + BY\bar{w} + BN_\dagger c, \quad \text{on } [a, \rho(b)]_{\mathbb{T}}, \tag{3.38}$$

$$M \begin{pmatrix} \bar{x}(a) \\ \bar{x}(b) \end{pmatrix} = d, \tag{3.39}$$

$$\bar{u} = Y\bar{w} + N_\dagger(c - K\bar{x}), \quad \text{on } [a, \rho(b)]_{\mathbb{T}}, \tag{3.40}$$

where $\bar{w}(\cdot)$ is in $\mathrm{C_{prd}}$.

Let

$$\tilde{C}(t) := [I + \mu(A(t) - B(t)N_{\dagger}(t)K(t))]^{-1}, \quad t \in [a, \rho(b)]_{\mathbb{T}}, \tag{3.41}$$

and define the $(m-k) \times (m-k)$-matrix function S on $[a, \rho(b)]_{\mathbb{T}}$ by

$$S(t) := Y^T(t)\left[R - \mu B^T \tilde{C}^T (Q - K^T N_{\dagger}^T R)\right](t) Y(t). \tag{3.42}$$

When S is invertible, define the non-homogeneous *symplectic boundary value problem* by

$$(\underline{\mathbb{S}}) \qquad \begin{cases} \bar{x}^{\Delta} = \underline{\mathbb{A}}(t)\bar{x} + \underline{\mathbb{B}}(t)\bar{p} + B(t)\underline{\zeta}(t) \\ \bar{p}^{\Delta} = \underline{\mathbb{C}}(t)\bar{x} + \underline{\mathbb{D}}(t)\bar{p} + \underline{\rho}(t) \end{cases}$$

$$(\text{bdry}) \qquad M\begin{pmatrix} \bar{x}(a) \\ \bar{x}(b) \end{pmatrix} = d, \quad \begin{pmatrix} \bar{p}(a) \\ -\bar{p}(b) \end{pmatrix} = \Gamma \begin{pmatrix} \bar{x}(a) \\ \bar{x}(b) \end{pmatrix} + \begin{pmatrix} c_a \\ c_b \end{pmatrix} + M^T \bar{\gamma},$$

where
(3.43)

$$\begin{cases} \underline{\zeta} := \mu Y S^{-1} Y^T B^T \tilde{C}^T [(Q - K^T N_{\dagger}^T R) N_{\dagger} c + z - K^T N_{\dagger}^T k] + N_{\dagger} c, \\ \underline{\rho} := \tilde{C}^T [\mu (Q - K^T N_{\dagger}^T R) Y S^{-1} Y^T B^T \tilde{C}^T + I][(Q - K^T N_{\dagger}^T R) N_{\dagger} c + z - K^T N_{\dagger}^T k], \\ \underline{\mathbb{A}} := A - B N_{\dagger} K + B Y S^{-1} Y^T \mathcal{V}, \\ \underline{\mathbb{B}} := B Y S^{-1} Y^T B^T \tilde{C}^T, \\ \underline{\mathbb{C}} := \tilde{C}^T [P - K^T N_{\dagger}^T Q^T + (Q - K^T N_{\dagger}^T R)(Y S^{-1} Y^T \mathcal{V} - N_{\dagger} K)], \\ \underline{\mathbb{D}} := \tilde{C}^T \left[K^T N_{\dagger}^T B^T - A^T + (Q - K^T N_{\dagger}^T R) Y S^{-1} Y^T B^T \tilde{C}^T\right], \end{cases}$$

and

$$\underline{\mathcal{V}} := \mu B^T \tilde{C}^T (P - Q N_{\dagger} K) - (I + \mu B^T \tilde{C}^T K^T N_{\dagger}^T)(Q^T - R N_{\dagger} K). \tag{3.44}$$

When F is *purely* quadratic (i.e., $c_a = c_b = 0$, $z \equiv 0$ and $k \equiv 0$), then $\zeta \equiv 0$ and $\rho \equiv 0$, and hence system $(\underline{\mathbb{S}})$ become *homogeneous*.

In the continuous time setting (i.e., $\mu \equiv 0$), and F is *purely* linear-quadratic, system $(\underline{\mathbb{S}})$ and (bdry) reduce to the homogeneous *Hamiltonian* system (5.8)–(5.11) in [34].

In the special case where the mixed state-control constraints are absent (i.e., $K \equiv 0$, $N \equiv 0$, and $c \equiv 0$), no endpoint constraints are present (i.e., $M = 0$ and $d = 0$), and F is *purely* quadratic, the coefficients (3.43) and system $(\underline{\mathbb{S}})$ reduce to the coefficients and their corresponding homogeneous symplectic system obtained in [24, Theorem 4.8]. Calculations prove that the coefficients $\underline{\mathbb{A}}, \underline{\mathbb{B}}, \underline{\mathbb{C}}$, and $\underline{\mathbb{D}}$ in (3.43) actually define a time scale *symplectic system.*

Proposition 3.17 (Weak-normality and symplectic systems for $(\mathcal{LQ})$)**.** *Assume that $[I + \mu(A - BN_{\dagger}K)](t)$ and $S(t)$ are invertible for all t in $[a, \rho(b)]$. Then, the following conditions are equivalent.*

(a) *The problem* $(\mathcal{LQ})$ *is weakly-normal at the extremal* $(\bar{x}, \bar{u})$ *with associated* $\bar{\Lambda} := (1, \bar{\lambda}, \bar{p}, \bar{\gamma})$.

(b) $(\bar{x}, \bar{p}, \bar{\gamma})$ *solve the symplectic boundary value problem* $(\underline{\mathbb{S}})$ *and* (*bdry*), *and*

$$\bar{u} = (YS^{-1}Y^T\underline{\mathcal{V}} - N_\dagger K)\bar{x} + YS^{-1}Y^T B^T \tilde{C}^T \bar{p} + \underline{\zeta}, \tag{3.45}$$

where $\underline{\zeta}$ *and* $\underline{\mathcal{V}}$ *are given in* (3.43) *and in* (3.44), *respectively.*

Proof. The problem $(\mathcal{LQ})$ is weakly normal at the extremal $(\bar{x}, \bar{u})$ means that there exist $\Lambda := (\lambda_0 = 1, \bar{\lambda}, \bar{p}, \bar{\gamma})$ satisfying (3.8)–(3.10) and there exists $\bar{w}$ such that $(\bar{x}, \bar{u})$ satisfy (3.38)–(3.40). Equation (3.9) where $\lambda_0 = 1$, is equivalent to

$$\bar{\lambda} = N_\dagger^T(B^T\bar{p}^\sigma - Q^T\bar{x} - R\bar{u} - k), \tag{3.46}$$

$$Y^T(-B^T\bar{p}^\sigma + Q^T\bar{x} + R\bar{u} + k) = 0. \tag{3.47}$$

Use (3.46), $\bar{p}^\sigma = \bar{p} + \mu\bar{p}^\Delta$, and (3.40) into (3.8) and solve for $\bar{p}^\Delta$ we get

$$\begin{aligned} \bar{p}^\Delta = \tilde{C}^T\Big[&(K^TN_\dagger^TB^T - A^T)\bar{p} + [P - K^TN_\dagger^TQ^T - (Q - K^TN_\dagger^TR)N_\dagger K]\bar{x} \\ &+ (Q - K^TN_\dagger^TR)(Y\bar{w} + N_\dagger c) + z - K^TN_\dagger^Tk\Big], \end{aligned} \tag{3.48}$$

$$\begin{aligned} \bar{p}^\sigma = \tilde{C}^T\Big[&\bar{p} + \mu[P - K^TN_\dagger^TQ^T - (Q - K^TN_\dagger^TR)N_\dagger K]\bar{x} \\ &+ \mu(Q - K^TN_\dagger^TR)(Y\bar{w} + N_\dagger c) + \mu(z - K^TN_\dagger^T)\Big]. \end{aligned} \tag{3.49}$$

Substitute (3.49) and (3.40) into (3.47) and use that S is invertible to solve for $\bar{w}$, we obtain

$$\bar{w} = S^{-1}Y^T\Big[B^T\tilde{C}^T\Big(\bar{p} + \mu(Q - K^TN_\dagger^TR)N_\dagger c + \mu(z - K^TN_\dagger^Tk)\Big) + \underline{\mathcal{V}}\bar{x}\Big], \tag{3.50}$$

where $\underline{\mathcal{V}}$ is defined by (3.44). This also implies that $\bar{w}$ is in $\mathrm{C_{prd}}$.

Finally plug the equation of $\bar{w}$ from (3.50) into (3.48), (3.38) and (3.40) we obtain that $(\bar{x}, \bar{p})$ satisfy the symplectic system $(\underline{\mathbb{S}})$ and (bdry), where the coefficients are given by (3.43), and $\bar{u}$ satisfies (3.45). Therefore, Condition (b) holds true.

Assume that (b) holds, i.e., $(\bar{x}, \bar{p}, \bar{\gamma})$ solve $(\underline{\mathbb{S}})$ and (bdry). Let $\bar{u}$ defined by (3.45) and $\bar{w}$ by (3.50). Easy calculations show that $(\underline{\mathbb{S}})$(i) implies $(\bar{x}, \bar{w})$ satisfies (3.38) and (3.39), and $\bar{u}$ satisfies (3.40). Thus, $(\bar{x}, \bar{u})$ is feasible. Furthermore, the definition of $\bar{w}$ in (3.50) and $\bar{u}$ gives that (3.47) holds true, and hence, there must be a function $\bar{\lambda}$ such that

$$-B^T\bar{p}^\sigma + Q^T\bar{x} + R\bar{u} + k = -N^T\bar{\lambda},$$

which yields that $\bar{\lambda}$ satisfies (3.46) and whence, $\bar{\lambda}$ is in $\mathrm{C_{prd}}$ and (3.9) holds for $\lambda_0 = 1$.

On the other hand, from $(\underline{\mathbb{S}})$(ii) we obtain that (3.32) holds true for $\lambda_0 = 1$ and $\bar{\lambda}$ defined by (3.46). Therefore, $(\mathcal{LQ})$ is weakly normal at $(\bar{x}, \bar{u})$, and condition (a) is satisfied. ∎

Using Proposition 3.17 we can re-state Corollary 3.11 in terms of the symplectic system ($\underline{\mathbb{S}}$).

Corollary 3.18 (Characterization of optimality and symplectic system in ($\mathcal{LQ}$))**.** *Assume that* $[I+\mu(A-BN_{\dagger}K)](t)$ *and* $S(t)$*, defined by* (3.42)*, are invertible for all* t *in* $[a,\rho(b)]_{\mathbb{T}}$*. Then, the following conditions are equivalent.*

(a) $(\bar{x},\bar{u})$ *is a (weak local or global) optimal for* ($\mathcal{LQ}$).
(b) *There exist a function* $\bar{p}$ *and a vector* $\bar{\gamma}$ *such that* $(\bar{x},\bar{p},\bar{\gamma})$ *solve* ($\underline{\mathbb{S}}$) *with (bdry),* $\bar{u}$ *satisfies* (3.45)*, and* $F''(\bar{x},\bar{u};\eta,v)\geq 0$ *for all admissible pairs.*

Remark 3.19. When F is purely linear-quadratic like F'' is, then $2F(\eta,v)=F''(\bar{x},\bar{u};\eta,v)$ and in this case, equations (3.8)–(3.10) together with equations (1.7)–(1.9) are called the *Jacobi system* (**J**) for ($\mathcal{LQ}$). The question of characterizing the non-negativity and positivity of $F(\eta,v)$ over the admissible pairs in terms of the Jacobi system (**J**) remains open for the general case of time scales.

To the symplectic system ($\underline{\mathbb{S}}$) and its boundary conditions (bdry) obtained from ($\mathcal{LQ}$) there correspond a quadratic *diagonal functional* $\underline{\mathbb{F}}(x,u)$ and a *diagonal* linear-quadratic problem ($\underline{\mathbf{D}}$)defined as:

$$\text{minimize}\quad \underline{\mathbb{F}}(x,p):=c_a^T x(a)+c_b^T x(b)+\frac{1}{2}\begin{pmatrix}x(a)\\x(b)\end{pmatrix}^T\Gamma\begin{pmatrix}x(a)\\x(b)\end{pmatrix}+\int_a^b\Big[\underline{z}^T x+\underline{k}^T p$$

$$(\underline{\mathbf{D}})\qquad +\frac{1}{2}\Big(x^T\underline{\mathbb{C}}^T(I+\mu\underline{\mathbb{A}})x+2\mu x^T\underline{\mathbb{C}}^T\underline{\mathbb{B}}\,p+p^T(I+\mu\underline{\mathbb{D}}^T)\underline{\mathbb{B}}p\Big)\Big](t)\Delta t$$

subject to: $x\in \mathrm{C}^1_{\mathrm{prd}}[a,b]_{\mathbb{T}}$ and $u\in \mathrm{C}_{\mathrm{prd}}[a,\rho(b)]_{\mathbb{T}}$ satisfying

$$x^{\Delta}(t)=\underline{\mathbb{A}}(t)x(t)+\underline{\mathbb{B}}(t)p(t)+B(t)\zeta(t)\quad t\in[a,\rho(b)]_{\mathbb{T}},\tag{3.51}$$

$$M\begin{pmatrix}x(a)\\x(b)\end{pmatrix}=d,\tag{3.52}$$

where

$$\underline{z}:=z+(\underline{\mathcal{V}}^T YS^{-T}Y^T-K^TN_{\dagger}^T)k\quad \underline{k}:=\tilde{C}BYS^{-T}Y^Tk,\tag{3.53}$$

and the coefficient $\underline{\mathbb{A}}$, $\underline{\mathbb{B}}$, $\underline{\mathbb{C}}$, $\underline{\mathbb{D}}$, and $\underline{\zeta}$ and $\underline{\mathcal{V}}$ are defined in (3.43) and (3.44).

Remark 3.20. Parallel to [24, Proposition 4.12], if (x,p) is feasible for ($\underline{\mathbf{D}}$), then for u defined through (x,p) via (3.45), we have (x,u) feasible for ($\mathcal{LQ}$) and $F(x,u)=\underline{\mathbb{F}}(x,p)+k^T\underline{\zeta}$. This yields that if $(\bar{x},\bar{u})$ is optimal for ($\mathcal{LQ}$), then $(\bar{x},\bar{p})$ is optimal for ($\underline{\mathbf{D}}$) where $\bar{p}$ is the function obtained from part(b) of Corollary 3.18. However, the converse of this statement is not always true, due to the fact that the feasibility of (x,u) in ($\mathcal{LQ}$) does not in general produce a function p with (x,p) feasible in ($\underline{\mathbf{D}}$). On the other hand, the assumption parallel to [24, Proposition 4.12], namely, for all t, $(Y^TB)(t)$ is of full rank), guarantees the converse statement, but again this assumption is too strong to impose on control problems. Therefore, obtaining the optimality for $(\mathcal{LQ})$ from that of $(\underline{\mathbf{D}})$ remains an open problem.

4. SYMPLECTIC QUADRATIC FORMS

Set

$$\begin{aligned}\mathcal{F}(x,q) := \frac{1}{2}\begin{pmatrix} x(a) \\ x(b) \end{pmatrix}^T \Gamma \begin{pmatrix} x(a) \\ x(b) \end{pmatrix} + \frac{1}{2}\int_a^b [x^T\mathcal{C}^T(I+\mu\mathcal{A})x \\ + 2\mu x^T\mathcal{C}^T\mathcal{B}q + q^T(I+\mu\mathcal{D})^T\mathcal{B}q](t)\Delta t.\end{aligned} \tag{4.1}$$

The *symplectic quadratic problem*, $(\mathcal{P})$, is defined as:

$$(\mathcal{P}) \qquad \text{minimize} \quad \mathcal{F}(x,q)$$

subject to: $x \in \mathrm{C}^1_{\mathrm{prd}}[a,b]_{\mathbb{T}}$ and $q \in \mathrm{C}_{\mathrm{prd}}[a,\rho(b)]_{\mathbb{T}}$ satisfying

$$x^\Delta(t) = \mathcal{A}(t)x(t) + \mathcal{B}(t)q(t), \quad t \in [a,\rho(b)]_{\mathbb{T}}, \tag{4.2}$$

$$\mathcal{M}\begin{pmatrix} x(a) \\ x(b) \end{pmatrix} = d, \tag{4.3}$$

where $\mathcal{M}$ is an $r \times 2n$-matrix, Γ is $2n \times 2n$-matrix, and $\mathcal{A}, \mathcal{B}, \mathcal{C}$, and $\mathcal{D}$ are in $\mathrm{C}_{\mathrm{prd}}[a,\rho(b)]_{\mathbb{T}}$ real $n \times n$-matrix functions such that the $2n \times 2n$-matrix function $\mathcal{S} := \begin{pmatrix} \mathcal{A} & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{pmatrix}$ satisfies the identity

$$\mathcal{S}^T(t)\mathcal{J} + \mathcal{J}\mathcal{S}(t) + \mu(t)\mathcal{S}^T(t)\mathcal{J}\mathcal{S}(t) = 0 \quad t \in [a,\rho(b)]_{\mathbb{T}}, \tag{4.4}$$

where $\mathcal{J} := \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}$ is skew-symmetric matrix. The symplecticity equation (4.4) translates to having (2.45) and (2.46) hold, where $\mathbb{A}, \mathbb{B}, \mathbb{C}$, and $\mathbb{D}$ are now $\mathcal{A}, \mathcal{B}, \mathcal{C}$ and $\mathcal{D}$, respectively.

The symplectic problem $(\mathcal{P})$ is a special form of the linear -quadratic control problem *without* shift in x, $(\mathcal{LQ})$, where, Γ and d are the same, and

$$\begin{cases} A := \mathcal{A}; \quad B := \mathcal{B}; \quad P := \mathcal{C}^T(I+\mu\mathcal{A}); \quad Q := \mu\mathcal{C}^T\mathcal{B}; \quad R := (I+\mu\mathcal{D})^T\mathcal{B}; \\ m = n; \quad u := q; \quad K \equiv N \equiv 0; \quad Y \equiv I; \quad \tilde{C} = \tilde{\mathcal{A}} := (I+\mu\mathcal{A})^{-1}; \\ c \equiv 0; \quad c_a = c_b = 0; \quad z \equiv 0; \quad k \equiv 0. \end{cases} \tag{4.5}$$

Given that $\mathcal{A}, \mathcal{B}, \mathcal{C}$, and $\mathcal{D}$ satisfy (2.45) yields that $P(t)$ and $R(t)$ in (4.5) are symmetric.

Similarly to Section 3, (x,q) is *feasible* (respectively, *admissible*) for $(\mathcal{P})$ means that (4.2) and (4.3) (respectively, (4.2) and $\mathcal{M}\begin{pmatrix} x(a) \\ x(b) \end{pmatrix} = 0$) hold.

Remark 4.1. If (x, q) is feasible for $(\mathcal{P})$ and if for some $p \in \mathrm{C}_{\mathrm{prd}}$ we have $\mathcal{B}q = \mathcal{B}p$ on $[a, \rho(b)]_{\mathbb{T}}$, then (x, p) is also feasible for $(\mathcal{P})$ and $\mathcal{F}(x, q) = \mathcal{F}(x, p)$. In fact, since $R := (I + \mu\mathcal{D})^T\mathcal{B}$ is symmetric, then by writing $\mathcal{B} = \mathcal{B}\mathcal{B}^\dagger\mathcal{B}$, where $\mathcal{B}^\dagger$ is the Penrose inverse of $\mathcal{B}$, we have $q^TRq = q^T\mathcal{B}^T(I + \mu\mathcal{D})\mathcal{B}^\dagger\mathcal{B}q$. Thus, if $\mathcal{B}q = \mathcal{B}p$ for $p, q \in \mathrm{C}_{\mathrm{prd}}$, we obtain that $q^TRq = p^T\mathcal{B}^T(I + \mu\mathcal{D})\mathcal{B}^\dagger\mathcal{B}p = p^TRp$.

Using the coefficients (4.5) into $(\mathcal{L})$ and $(\mathcal{Q})$, the first and second variations of the problem $(\mathcal{P})$ at a feasible pair $(\bar{x}, \bar{q})$ in the direction of an admissible pair (η, v) reduce to:

$$\mathcal{F}'(\bar{x}, \bar{q}; \eta, v) := \begin{pmatrix} \bar{x}(a) \\ \bar{x}(b) \end{pmatrix}^T \Gamma \begin{pmatrix} \eta(a) \\ \eta(b) \end{pmatrix} + \int_a^b \Big[\bar{x}^T\mathcal{C}^T(I + \mu\mathcal{A})\eta + \bar{q}^T\mu\mathcal{B}^T\mathcal{C}\eta$$

$$(\mathscr{L}) \qquad + \bar{x}^T\, \mu\mathcal{C}^T\mathcal{B}\, v + \bar{q}^T\, (I + \mu\mathcal{D})^T\mathcal{B}\, v\Big](t)\, \Delta t$$

$$\mathcal{F}''(\bar{x}, \bar{q}; \eta, v) := 2\mathcal{F}(\eta, v) := \begin{pmatrix} \eta(a) \\ eta(b) \end{pmatrix}^T \Gamma \begin{pmatrix} \eta(a) \\ \eta(b) \end{pmatrix} + \int_a^b \Big[\eta^T\, \mathcal{C}^T(I + \mu\mathcal{A})\eta$$

$$(\mathscr{Q}) \qquad + 2\eta^T\mu\mathcal{C}^T\mathcal{B}v + v^T(I + \mu\mathcal{D})^T\mathcal{B}v\Big](t)\Delta t.$$

Note that $\mathcal{F}'(\bar{x}, \cdot; \eta, v)$ is invariant over $\{q \in \mathrm{C}_{\mathrm{prd}} : \mathcal{B}q = \mathcal{B}\bar{q}$, on $[a, \rho(b)]_{\mathbb{T}}\}$, and $\mathcal{F}''$ is independent of $(\bar{x}, \bar{q})$.

An immediate consequence of Proposition 3.2 is the following characterization of optimality for $(\mathcal{P})$ in terms of $\mathcal{F}'$ and $\mathcal{F}''$, which is $2\mathcal{F}$.

Proposition 4.2 (First and second variations for $(\mathcal{P})$). *Let $(\bar{x}, \bar{q})$ be feasible for $(\mathcal{P})$. The following statements are equivalent.*

(a) $(\bar{x}, \bar{q})$ *is a weak local minimum for* $(\mathcal{P})$.
(b) *For all admissible pairs* (η, v) *we have* $\mathcal{F}'(\bar{x}, \bar{q}; \eta, v) = 0$ *and* $\mathcal{F}''(\bar{x}, \bar{q}; \eta, v) \geq 0$.
(c) $(\bar{x}, \bar{q})$ *is a global minimum for* $(\mathcal{P})$.

Remark 4.3. In part (ii) of [25, Proposition 4.3], it is shown that when $d = 0$ and under both the invertibility of $I + \mu\mathcal{A}$ and the $\mathcal{M}$-controllability of the equation of motion (which are required to apply therein [23, Theorems 9.6 and 9.7]), the optimality in $(\mathcal{P})$ implies that $\mathcal{F}'(\bar{x}, \bar{q}; \eta, v) = 0$ and $\mathcal{F}''(\bar{x}, \bar{q}; \eta, v) = 2\mathcal{F}(\eta, v) \geq 0$ for all admissible pairs (η, v). Therefore, Proposition 4.2 implies for this special case that *neither* the invertibility of $I + \mu\mathcal{A}$ *nor* the $\mathcal{M}$-controllability of the equation of motion are needed in order that the statement in part (ii) of [25, Proposition 4.3] be valid, and furthermore, its converse is also true.

When $d = 0$, there is an extensive work in the literature to characterize the non-negativity of $\mathcal{F}$ in terms of the natural conjoint basis or the principal solution of the corresponding symplectic system, or in terms of an associated implicit Riccati equation (see e.g., [25, Theorem 3.2], [21, Theorems 5.1 and 5.3] and [22, Theorem 4.2]). Each of these equivalent conditions is easier to verify than the non-negativity of the $\mathcal{F}$ over all possible admissible pairs.

To characterize the first order condition in Proposition 4.2, we shall use the weak normality notion and the corresponding *homogeneous* symplectic boundary value problem associated with $(\mathcal{P})$, for being easier conditions to verify.

Using the coefficients from (4.5) into Definition 3.3 where $\lambda_0 = 1$, we obtain the following definition of *weak normality* of $(\mathcal{P})$ at a feasible pair $(\bar{x}, \bar{q})$.

Definition 4.4 (Weak normality of $(\mathcal{P})$)**.** The problem $(\mathcal{P})$ is *weakly normal* at a feasible pair $(\bar{x}, \bar{q})$ means that there exist $\bar{p}$ in $\mathrm{C}^1_{\mathrm{prd}}[a,b]_{\mathbb{T}}$ and $\bar{\gamma}$ in $\mathbb{R}^k$ such that:

$$\begin{cases} \bar{x}^{\Delta} = \mathcal{A}\bar{x} + \mathcal{B}\bar{q}, \quad \text{on } [a,\rho(b)]_{\mathbb{T}}, \\ \bar{p}^{\Delta} = -\mathcal{A}^T\bar{p}^{\sigma} + \mathcal{C}^T(I+\mu\mathcal{A})\bar{x} + \mu\mathcal{C}^T\mathcal{B}\bar{q}, \quad \text{on } [a,\rho(b)]_{\mathbb{T}}, \\ -\mathcal{B}^T\bar{p}^{\sigma} + \mu\mathcal{B}^T\mathcal{C}\bar{x} + (I+\mu\mathcal{D}^T)\mathcal{B}\bar{q} = 0, \quad \text{on } [a,\rho(b)]_{\mathbb{T}}, \\ \mathcal{M}\begin{pmatrix}\bar{x}(a)\\ \bar{x}(b)\end{pmatrix} = d, \qquad \begin{pmatrix}\bar{p}(a)\\ -\bar{p}(b)\end{pmatrix} = \Gamma\begin{pmatrix}\bar{x}(a)\\ \bar{x}(b)\end{pmatrix} + \mathcal{M}^T\bar{\gamma}. \end{cases} \tag{4.6}$$

Remark 4.5. Note that Definition 4.4 yields that if $(\mathcal{P})$ is *weakly normal* at $(\bar{x}, \bar{q})$ with a multipliers $(\bar{p}, \bar{\gamma})$ and if for some $q \in \mathrm{C}_{\mathrm{prd}}$ we have $\mathcal{B}q = \mathcal{B}\bar{q}$ on $[a,\rho(b)]_{\mathbb{T}}$ then $(\bar{x}, q)$ is feasible and $(\mathcal{P})$ is *weakly normal* at $(\bar{x}, q)$ with the same multipliers $(\bar{p}, \bar{\gamma})$.

If we intend to apply to problem $(\mathcal{P})$ the results of Section 3 pertaining symplectic systems, we would define the *symplectic boundary value problem* corresponding to $(\mathcal{P})$ in the same way we defined it for $(\mathcal{LQ})$, that is, by using the coefficients given by (3.43), which require, in addition to the invertibility of $I + \mu\mathcal{A}$, the invertibility of $S(t)$ defined via (3.42). From (4.5) and (2.45), it follows that here $S = \mathcal{B}^T\tilde{\mathcal{A}}^T$, and hence, the invertibility of $S(t)$ is equivalent to the invertibility of $\mathcal{B}(t)$. This assumption renders the problem $(\mathcal{P})$ a simple calculus of variations problem, which is too restrictive for this study. Thus, in order to obtain general results for $(\mathcal{P})$ we should not apply the results known for $(\mathcal{LQ})$ concerning symplectic systems, but instead, we must tackle the problem $(\mathcal{P})$ directly.

Nevertheless, if for the time being the invertibility of $(I + \mu\mathcal{A})(t)$ and $\mathcal{B}(t)$ is assumed, we notice when calculating the coefficients of the symplectic system associated to $(\mathcal{P})$ via (3.43), that neither $\tilde{\mathcal{A}}$ nor S^{-1} appears in these coefficients. In fact, we obtain $\underline{\mathbb{A}} = \mathcal{A}, \underline{\mathbb{B}} = \mathcal{B}, \underline{\mathbb{C}} = \mathcal{C}, \underline{\mathbb{D}} = \mathcal{D}, c_a = c_b = 0, \underline{z} \equiv 0$, and $\underline{k} \equiv 0$; which is not surprising given that $(\mathcal{P})$ itself is *symplectic.* Therefore, *without* any invertibility assumption, the symplectic boundary value problem corresponding to $(\mathcal{P})$ is now *homogeneous* and is defined through the coefficients $\mathcal{A}, \mathcal{B}, \mathcal{C}$, and $\mathcal{D}$, that is,

$$(\mathscr{S}) \qquad \begin{cases} \bar{x}^{\Delta} = \mathcal{A}(t)\bar{x} + \mathcal{B}(t)\bar{p} & t \in [a, \rho(b)]_{\mathbb{T}}, \\ \bar{p}^{\Delta} = \mathcal{C}(t)\bar{x} + \mathcal{D}(t)\bar{p} & t \in [a, \rho(b)]_{\mathbb{T}}, \end{cases}$$

$$(bdry) \qquad \mathcal{M}\begin{pmatrix} \bar{x}(a) \\ \bar{x}(b) \end{pmatrix} = d, \quad \begin{pmatrix} \bar{p}(a) \\ -\bar{p}(b) \end{pmatrix} = \Gamma \begin{pmatrix} \bar{x}(a) \\ \bar{x}(b) \end{pmatrix} + \mathcal{M}^T \bar{\gamma}.$$

Note that the system $(\mathscr{S})$ is symplectic because the matrix of coefficients $\mathcal{S} := \begin{pmatrix} \mathcal{A} & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{pmatrix}$ satisfies (4.4). Therefore, $(\bar{x}, \bar{p})$ solves $(\mathscr{S})$ if and only if it solves the time-reversed system $(\mathscr{S}^-)$

$$(\mathscr{S}^-) \qquad \begin{cases} \bar{x}^{\Delta} = -\mathcal{D}^T(t)\bar{x}^{\sigma} + \mathcal{B}^T(t)\bar{p}^{\sigma} & t \in [a, \rho(b)]_{\mathbb{T}}, \\ \bar{p}^{\Delta} = \mathcal{C}^T(t)\bar{x}^{\sigma} - \mathcal{A}^T(t)\bar{p}^{\sigma} & t \in [a, \rho(b)]_{\mathbb{T}}. \end{cases}$$

Unlike $(\mathcal{LQ})$, Problem $(\mathcal{P})$, being *symplectic*, has a special structure that stems from the fact that its associated coefficients (P, Q, R, A, B) defined through (4.5) are inter-dependent in a specific manner via (4.4). Consequently, unlike Proposition 3.17, we shall show directly that *without* any extra assumption on $I + \mu\mathcal{A}$ or on $\mathcal{B}$, the weak normality of $(\mathcal{P})$ at $(\bar{x}, \bar{q})$ with associated multipliers $(\bar{p}, \bar{\gamma})$ is equivalent not only to $(\bar{x}, \bar{p}, \bar{\gamma})$ solving the *symplectic boundary value problem* with $\mathcal{B}\bar{p} = \mathcal{B}\bar{q}$, but also to the weak normality of $(\mathcal{P})$ at $(\bar{x}, \bar{p})$ whose associated multipliers are also $(\bar{p}, \bar{\gamma})$ and such that $\mathcal{B}\bar{p} = \mathcal{B}\bar{q}$.

Proposition 4.6 (Weak-normality and symplectic systems for $(\mathcal{P})$)**.** *The following conditions are equivalent.*

- (a) *Problem $(\mathcal{P})$ is weakly-normal at $(\bar{x}, \bar{q})$, with multipliers $(\bar{p}, \bar{\gamma})$ satisfying* (4.6)*.*
- (b) *$(\bar{x}, \bar{p}, \bar{\gamma})$ solve the homogeneous symplectic boundary value problem $(\mathscr{S})$ and (bdry), and $\mathcal{B}(t)\bar{q}(t) = \mathcal{B}(t)\bar{p}(t)$, for all $t \in [a, \rho(b)]_{\mathbb{T}}$.*
- (c) *The problem $(\mathcal{P})$ is weakly-normal at $(\bar{x}, \bar{p})$, with multipliers $(\bar{p}, \bar{\gamma})$ satisfying* (4.6)*, and $\mathcal{B}(t)\bar{q}(t) = \mathcal{B}(t)\bar{p}(t)$, for all $t \in [a, \rho(b)]_{\mathbb{T}}$.*

Remark 4.7. As a consequence of the equivalence between (b) and (c) in Proposition 4.6 and by using therein $\bar{q} := \bar{p}$ at which $\mathcal{B}\bar{q} = \mathcal{B}\bar{p}$ is trivally satisfied, we obtain the following equivalence:

$(\bar{x}, \bar{p}, \bar{\gamma})$ solve $(\mathscr{S})$ and $(bdry)$ $\iff$ $(\mathcal{P})$ is weakly-normal at $(\bar{x}, \bar{p})$, with multiplier $(\bar{p}, \bar{\gamma})$ satisfying (4.6).

Proof of Proposition 4.6*.*

$(a) \Rightarrow (b)$: Let (a) hold and let $(\bar{x}, \bar{q}, \bar{p}, \bar{\gamma})$ satisfy (4.6). Use the equation of motion (4.6)(i) into (4.6)(ii), we get

$$\bar{p}^{\Delta} = \mathcal{C}^T(t)\bar{x}^{\sigma} - \mathcal{A}^T(t)\bar{p}^{\sigma}. \tag{4.7}$$

Use from (2.45) that $\mu\mathcal{B}^T\mathcal{C} = \mathcal{A} + \mathcal{D}^T + \mu\mathcal{D}^T\mathcal{A}$ into (4.6)(iii) and then use (4.6)(i) to obtain

$$\mathcal{B}^T\bar{p}^\sigma = \mathcal{D}^T\bar{x} + (I + \mu\mathcal{D}^T)\bar{x}^\Delta = \bar{x}^\Delta + \mathcal{D}^T\bar{x}^\sigma.$$

Thus, the pair $(\bar{x}, \bar{p})$ solves $(\mathscr{S}^-)$, and hence, it solves $(\mathscr{S})$. Therefore, $(\bar{x}, \bar{p}, \bar{\gamma})$ solves the *symplectic boundary value problem*, $(\mathscr{S})$ and (*bdry*). Now, given that both $(\bar{x}, \bar{q})$ and $(\bar{x}, \bar{p})$ satisfy the equation of motion, we obtain that $\mathcal{B}\bar{q} = \mathcal{B}\bar{p}$, that is, (b) holds.

(*b*) $\Rightarrow$ (*c*): Suppose that $(\bar{x}, \bar{p}, \bar{\gamma})$ solves $(\mathscr{S})$ and (bdry), thus, it solves $(\mathscr{S}^-)$ and (bdry) and,

$$\bar{x}^\Delta = \mathcal{A}\bar{x} + \mathcal{B}\bar{p}, \quad \bar{x}^\sigma = (I + \mu\mathcal{A})\bar{x} + \mu\mathcal{B}\bar{p}. \tag{4.8}$$

Replace the equation for $\bar{x}^\sigma$ from (4.8)(ii) into $(\mathscr{S}^-)$(ii), it follows that the adjoint equation,(4.6)(ii), is satisfies where $\bar{q} := \bar{p}$. On the other hand, from $(\mathscr{S}^-)$(i), in which we replace $\bar{x}^\Delta$ and $\bar{x}^\sigma$ from (4.8) and we use $\mathcal{A} + \mathcal{D}^T + \mu\mathcal{D}^T\mathcal{A} = \mu\mathcal{B}^T\mathcal{C}$, we obtain (4.6)(iii), where $\bar{q} = \bar{p}$. This shows that $(\bar{x}, \bar{p}, \bar{p}, \bar{\gamma})$ satisfy (4.6).

(*c*) $\Rightarrow$ (*a*) let $(\bar{x}, \bar{p}, \bar{p}, \bar{\gamma})$ satisfy (4.6) and $\mathcal{B}\bar{p} = \mathcal{B}\bar{q}$. Replacing $\mathcal{B}\bar{p}$ by $\mathcal{B}\bar{q}$ it followsm that $(\bar{x}, \bar{q}, \bar{p}, \bar{\gamma})$ satisfy (4.6). ■

Now, without assuming the invertibility of $\mathcal{B}$ nor the $\mathcal{M}$-controllability of the linear system (4.2), we can combine Theorem 3.10 applied to $(\mathcal{P})$ with Proposition 4.6 to readily obtain characterizions of the first order condition in Proposition 4.2 in terms of the weak-normality condition of $(\mathcal{P})$ and its equivalent condition involving the symplectic boundary value problem.

Theorem 4.8 (First variation, weakly normalilty, symplectic system of $(\mathcal{P})$)**.** *Assume that $I + \mu(t)\mathcal{A}(t)$ is invertible for all $t \in [a, \rho(b)]_{\mathbb{T}}$. Let $(\bar{x}, \bar{q})$ be feasible for $(\mathcal{P})$, then the following conditions are equivalent.*

(i) $F'(\bar{x}, \bar{q}; \eta, v) = 0$ *for all admissible pairs* (η, v).
(ii) *Problem* $(\mathcal{P})$ *is weakly normal at* $(\bar{x}, \bar{q})$, *with multipliers* $(\bar{p}, \bar{\gamma})$ *satisfying* (4.6).
(iii) $(\bar{x}, \bar{p}, \bar{\gamma}$ *solve the homogeneous symplectic boundary value problem* $(\mathscr{S})$ *and (bdry), and* $\mathcal{B}(t)\bar{q}(t) = \mathcal{B}(t)\bar{p}(t)$, *for all* $t \in [a, \rho(b)]_{\mathbb{T}}$.
(iv) *Problem* $(\mathcal{P})$ *is weakly-normal at* $(\bar{x}, \bar{p})$, *with multipliers* $(\bar{p}, \bar{\gamma})$ *satisfying* (4.6), *and* $\mathcal{B}(t)\bar{q}(t) = \mathcal{B}(t)\bar{p}(t)$, *for all* $t \in [a, \rho(b)]_{\mathbb{T}}$.

Remark 4.9. (a) The assumption that $I + \mu\mathcal{A}$ is invertible is not assumed in Proposition 4.6 and is *not* required to prove the implication (ii) $\Rightarrow$ (i).

(b) If we only assume the first part of (iii) holds, namely, $(\bar{x}, \bar{p}, \bar{\gamma})$ solve the symplectic boundary value problem $(\mathscr{S})$ and (bdry) or, its equivalent form in Remark 4.7, i.e., the first part of (iv), then easy calculations show that $\mathcal{F}'(\bar{x}, \bar{p}; \eta, v) = 0$, for all admissible pairs (η, v). In other words, without the $\mathcal{M}$-controllability of system (4.2) or the invertibility of $I + \mu\mathcal{A}$ we have:

$(\bar{x}, \bar{p}, \bar{\gamma})$ solve $(\mathscr{S})$ and (bdry) $\Rightarrow$ $(\bar{x}, \bar{p})$ is feasible and $\mathcal{F}'(\bar{x}, \bar{p}; \eta, v) = 0$ for all admissible pairs.

The converse of this implication is not true in general, since even under $I + \mu\mathcal{A}$ invertible when applying to $(\bar{x}, \bar{p})$ that (i) $\Rightarrow$ (ii), we would only get that there are multipliers p and γ such that $(\bar{x}, \bar{p}, p, \gamma)$ satisfies (4.6), and hence, from (ii) $\Rightarrow$ (iii), we have only that $(\bar{x}, p, \gamma)$ solve $(\mathscr{S})$ and (bdry). There is no reason that p be $\bar{p}$.

(c) Assume that $d = 0$, the $\mathcal{M}$-controllability of (4.2) holds, and the invertibility of $I + \mu\mathcal{A}$ is met. In part(i) of [25, Proposition 4.3] it is shown that whenever $\mathcal{F}(\eta, v) \geq 0$, for all feasible pairs, and a triplet $(\bar{x}, \bar{p}, \bar{\gamma})$ solves $(\mathscr{S})$ and (bdry), then $\mathcal{F}'(\bar{x}, \bar{p}; \eta, v) = 0$, for all admissible pairs, and $(\bar{x}, \bar{p})$ is optimal for $(\mathcal{P})$. However, we claim that this result is true *without* the M-controllability or the invertibility of $I + \mu\mathcal{A}$. Indeed, we first note that in this case admissible and feasible pairs are the same. In addition, by part (b) of this Remark, $(\bar{x}, \bar{p}, \bar{\gamma})$ solving $(\mathscr{S})$ and (bdry) yields that $\mathcal{F}'(\bar{x}, \bar{p}; \eta, v) = 0$ for all admissible or feasible pairs. Furthermore, using $\mathcal{F}(\bar{x}, \bar{p}) = \frac{1}{2}\mathcal{F}'(\bar{x}, \bar{p}; \bar{x}, \bar{p}) = 0$ and $\mathcal{F} \geq 0$, it follows that $(\bar{x}, \bar{p})$ must be *optimal* for $(\mathcal{P})$.

An immediate consequence of combining Theorem 4.8 with Proposition 4.2 is the following Corollary which is a characterization of the optimality for $(\mathcal{P})$ in terms of each of the corresponding weak normality or the symplectic boundary value problem, together with the non-negativity of the second variation.

Corollary 4.10 (Characterization of optimality, weak normality, symplectic system for $(\mathcal{P})$)**.** *Assume that $I+\mu(t)A(t)$ is invertible for all t in $[a, \rho(b)]$. Then, the following conditions are equivalent.*

(a) $(\bar{x}, \bar{q})$ *is (weak local or global) optimal for* $(\mathcal{P})$.
(b) *Problem* $(\mathcal{P})$ *is weakly normal at* $(\bar{x}, \bar{q})$ *with multipliers* $(\bar{p}, \bar{\gamma})$ *satisfying* (4.6)*, and* $\mathcal{F}''(\bar{x}, \bar{q}; \eta, v) \geq 0$ *for all admissible pairs.*
(c) $(\bar{x}, \bar{p}, \bar{\gamma})$ *solve the symplectic boundary value problem* $(\mathscr{S})$ *and (bdry), with* $\mathcal{B}(t)\bar{q}(t) = \mathcal{B}(t)\bar{p}(t)$*, for all* $t \in [a, \rho(b)]_{\mathbb{T}}$*, and* $\mathcal{F}''(\bar{x}, \bar{q}; \eta, v) \geq 0$ *for all admissible pairs.*
(d) *The problem* $(\mathcal{P})$ *is weakly-normal at* $(\bar{x}, \bar{p})$*, with multipliers* $(\bar{p}, \bar{\gamma})$ *satisfying* (4.6)*, with* $\mathcal{B}(t)\bar{q}(t) = \mathcal{B}(t)\bar{p}(t)$*, for all* $t \in [a, \rho(b)]_{\mathbb{T}}$*, and* $\mathcal{F}''(\bar{x}, \bar{q}; \eta, v) \geq 0$ *for all admissible pairs.*

Remark 4.11. (i) Let $(\bar{x}, \bar{q})$ be optimal for $(\mathcal{P})$ and $\bar{p}$ be its adjoint variable satisfying with $\bar{\gamma}$ condition (b) of Corollary 4.10. Then, $(\bar{x}, \bar{p})$ is also optimal for $(\mathcal{P})$ and the same multiplier pair $(\bar{p}, \bar{\gamma})$ satisfies with $(\bar{x}, \bar{p})$ (4.6). In fact, from (a) $\Rightarrow$ (c) in Corollary 4.10, it follows that $\mathcal{B}\bar{q} = \mathcal{B}\bar{p}$ and hence, by Remark 4.1, we have $(\bar{x}, \bar{p})$ is also optimal for $(\mathcal{P})$. Furthermore, the implication (a) $\Rightarrow$ (d) yields that $(\bar{p}, \bar{\gamma})$ is a multiplier corresponding to the optimal pair $(\bar{x}, \bar{p})$.

(ii) In [25, Proposition 4.3], when $d = 0$ and under both the invertibility of $I + \mu\mathcal{A}$ and the $\mathcal{M}$-controllability of the equation of motion, the implications (a) $\Rightarrow$ (b) and (a) $\Rightarrow$ (c) (without $\mathcal{B}\bar{q} = \mathcal{B}\bar{p}$) are shown to be true. Therefore, Corollary 4.10 is a generalization of [25, Proposition 4.3(ii)] to the case where the $\mathcal{M}$-controllability of the equation of motion does not hold.

Remark 4.12. (**General Remark**): All the results and the proofs in this paper remain valid if we replace $(x, u) \in \mathrm{C}^1_{\mathrm{prd}} \times \mathrm{C}_{\mathrm{prd}}$ by $(x, u) \in W^{1,2} \times L^2$. Of course, the dynamic equations in this setting would be satisfied almost everywhere. We refer to [32] and the references therein for the notions and properties of the Sobolev and Lebesgue spaces on time scales.

Acknowledgements

The author conducted a part of this project when she was supported by the National Science Foundation grant # DMS-0707789.

REFERENCES

[1] R. P. Agarwal, M. Bohner, P. J. Y. Wong, Sturm-Liouville eigenvalue problems on time scales, *Appl. Math. Comput.* **99** (1999), no. 2–3, 153–166.

[2] C. D. Ahlbrandt, M. Bohner, J. Ridenhour, Hamiltonian systems on time scales, *J. Math. Anal. Appl.* **250** (2000), no. 2, 561–578.

[3] C. D. Ahlbrandt, A. C. Peterson, *Discrete Hamiltonian Systems: Difference Equations, Continued Fractions, and Riccati Equations,* Kluwer Academic Publishers, Boston, 1996.

[4] F. M. Atici, D. C. Biles, A. Lebedinsky, An application of time scales to economics, *Math. Comput. Modelling* **43** (2006), no. 7–8, 718–726.

[5] M. Bohner, Calculus of variations on time scales, *Dynam. Systems Appl.* **13** (2004), no. 3–4, 339–349.

[6] M. Bohner, A. Peterson, *Dynamic Equations on Time Scales. An Introduction with Applications,* Birkhäuser, Boston, 2001.

[7] V. G. Boltyanskii, *Optimal Control of Discrete Systems,* John Wiley & Sons, New York - Toronto, 1978.

[8] O. Došlý, S. Hilger, A necessary and sufficient condition for oscillation of the Sturm-Liouville dynamic equation on time scales, In: "Dynamic Equations on Time Scales", R. P. Agarwal, M. Bohner, and D. O'Regan, editors. *J. Comput. Appl. Math.* **141** (2002), no. 1–2, 147–158.

[9] J. F. Bonnans, A. Hermant, Second-order analysis for optimal control problems with pure state constraints and mixed control-state constraints, *Ann. I. H. Poincar* **26** (2009), 561-598.

[10] L. Erbe, S. Hilger, Sturmian theory on measure chains, *Differential Equations Dynam. Systems* **1** (1993), no. 3, 223–244.

[11] L. Erbe, A. Peterson, Oscillation criteria for second order matrix dynamic equations on a time scale, In: "Dynamic Equations on Time Scales", R. P. Agarwal, M. Bohner, and D. O'Regan, editors. *J. Comput. Appl. Math.* **141** (2002), no. 1–2, 169–185.

[12] M. R. Hestenes, *Calculus of Variations and Optimal Control Theory,* John Wiley & Sons, New York, 1966.

[13] M. R. Hestenes, Quadratic variational theory, In: "Control Theory and the Calculus of Variations", pp. 1–37, A. V. Balakrishnan, editor. Academic Press, New York, 1969.

[14] R. Hilscher, V. Zeidan, Discrete optimal control: the accessory problem and necessary optimality conditions, *J. Math. Anal. Appl.* **243** (2000), no. 2, 429–452.

[15] R. Hilscher, V. Zeidan, Second order sufficiency criteria for a discrete optimal control problem, *J. Difference Equ. Appl.* **8** (2002), no. 6, 573–602.

[16] R. Hilscher, V. Zeidan, Symplectic difference systems: variable stepsize discretization and discrete quadratic functionals, *Linear Algebra Appl.* **367** (2003), 67–104.

[17] R. Hilscher, V. Zeidan, Calculus of variations on time scales: weak local piecewise $\mathrm{C}^1_{\mathrm{rd}}$ solutions with variable endpoints, *J. Math. Anal. Appl.* **289** (2004), no. 1, 143–166.

[18] R. Hilscher, V. Zeidan, Nonnegativity and positivity of a quadratic functional in the discrete calculus of variations: A survey, *J. Difference Equ. Appl.* **11** (2005), no. 9, 857–875.

[19] R. Hilscher, V. Zeidan, Legendre, Jacobi, and Riccati type conditions for time scale variational problem with application, *Dynam. Systems Appl.* **16** (2007), no. 3, 451–480.

[20] R. Hilscher, V. Zeidan, Time scale embedding theorem and coercivity of quadratic functionals, *Nonlinear Analysis (Munich)* **70** (2009), 3209–3226.

[21] R. Hilscher, V. Zeidan, Riccati equations for abnormal time scale quadratic functionals, *J. Differential Equations* **244** (2008), no. 6, 1410–1447.

[22] R. Hilscher, V. Zeidan, Applications of time scale symplectic systems without normality, *J. Math. Anal. Appl.* **340** (2008), no. 1, 451–465.

[23] R. Hilscher, V. Zeidan, Weak maximum principle and accessory problem for control problems on time scales, *Nonlinear Analysis* **70** (2009), 1–28.

[24] R. Hilscher, V. Zeidan, Symplectic structure of Jacobi systems on time scales, *Int. J. Difference Equ.* **5** (2010), no. 1, 55–81.

[25] R. Hilscher, V. Zeidan, Reid roundabout theorems for time scale symplectic systems, *Discrete dynamics and difference equations* World Sci. Publ., Hackensack, NJ, (2010), 267-288.

[26] W. G. Kelley, A. C. Peterson, *Difference Equations: An Introduction with Applications,* Academic Press, San Diego, 1991.

[27] F. L. Lewis, *Optimal Control,* John Wiley & Sons, Inc., New York, 1986.

[28] A. A. Milyutin, N. P. Osmolovskii, *Calculus of Variations and Optimal Control,* American Mathematical Society, Providence, RI, 1998.

[29] L. W. Neustadt, *Optimization. A Theory of Necessary Conditions, Princeton University Press, Princeton, NJ, 1976,* .

[30] N. P. Osmolovskii, Second-order conditions for a weak local minimum in an optimal control problem (necessity, sufficiency), *Soviet Math. Dokl.* **16** (1975), 1480–1484.

[31] W. T. Reid, *Ordinary Differential Equations,* Wiley, New York, 1971.

[32] B. P. Rynne, L^2 spaces and boundary value problems on time-scales, *J. Math. Anal. Appl.* **328** (2007), no. 2, 1217–1236.

[33] V. Zeidan, P. Zezza, The conjugate point condition for smooth control sets, *J. Math. Anal. Appl.* **132** (1988), no. 2, 572–589.

[34] V. Zeidan, The Riccati equation for optimal control problems with mixed state-control constraints:necessity and sufficiency, *SIAM J. Control and Optimization* **32** (1994), no. 5, 1297–1321.

[35] V. Zeidan, Admissible directions and generalized coupled points for optimal control problems, *Nonlinear Analysis: Theory, Methods and Applications* **26** (1996), no. 3, 479–507.